Fair Scheduling of Concurrent Transmissions in Directional Antenna Based WPANs/WLANs

L. Rajya Lakshmi and Biplab Sikdar
Department of Electrical and Computer Engineering
National University of Singapore, Singapore

***Abstract*—With their capability to support high data rates, millimeter-Wave (mmWave) communications are evolving as a promising and potential technology to support high data rate applications in short range networks. This paper addresses the problem of fair scheduling in mmWave wireless personal and local area networks (WPANs/WLANs) to support applications with varying quality of service (QoS) requirements. To ensure fairness while exploiting the spatial reuse facilitated by directional antennas, concurrent transmission scheduling in mmWave WPANs/WLANs is formulated as a multi-objective optimization problem. Two heuristic schedulers are developed to obtain a schedule in real-time. These schedulers first satisfy the minimum QoS requirements of as many flows as possible, and then, allocate the remaining bandwidth to various flows while ensuring long-term and short-term fairness among the flows. Results from extensive simulations conducted in a dense mmWave WPAN show that the proposed fair schedulers provide better fairness and throughput, compared to existing methods.**

I. Introduction

Experts from both industry and academia predict an explosive growth in mobile data rate requirements in the near future [1]. Communications in the millimeter wave (mmWave) band are evolving as a potential technology to address these rapidly increasing mobile data rate requirements. The mmWave band can be used as the wireless backhaul of heterogeneous cellular networks [2] and as an access technology in wireless personal and local area networks (WPANs/WLANs) [3]. With 7 GHz of unlicensed bandwidth available all over the world, mmWave communications in the 60 GHz band can support high data rate applications such as high speed Internet access, real-time streaming of high-definition television, etc.

Owing to high carrier frequency, path loss in the mmWave band is significantly higher compared to bands with lower carrier frequencies. This characteristic makes it necessary to use directional antennas both at the transmitters and the receivers to combat the high path loss and realize directional communications in the mmWave band [2], [3]. With the usage of directional antennas, multi-user interference reduces significantly and concurrent transmissions can take place to improve the network throughput. There exist various spatial time division multiple access (STDMA) techniques to exploit the spatial reuse in mmWave networks [2]–[5].

Fair scheduling has been an important issue in both wireline and wireless networks to provide weight based throughput guarantees to flows with different weights. However, the main focus of the existing STDMA scheduling methods is network throughput enhancement [2], [3], [5]. Also, there exist some methods that aim to maximize the number of quality of service (QoS) satisfied flows [4]. On the other hand, there exists considerable amount of research on fair scheduling in different types of networks [6], [7]. Most of these methods generate schedule one slot at a time and assume either perfect or partial per-flow knowledge. However, some of the unique characteristics of mmWave networks such as large number of slots per frame and very limited per-flow information, prevent the direct applicability of the existing fair scheduling methods in mmWave networks. Hence, in this paper, we first formulate the fair scheduling in mmWave networks as a multi-objective optimization problem, and then, develop two heuristic fair STDMA schedulers that generate near-optimal fair schedules.

The main contributions of this paper are three fold:

- To maximize the network throughput while ensuring fairness among the flows, concurrent transmission scheduling in directional antenna based networks is formulated as a multi-objective optimization problem.
- Two heuristic concurrent transmission schedulers that ensure short-term and long-term fairness among the flows are developed.
- Extensive simulations are conduced in a dense mmWave WPAN to demonstrate the fair allocation and QoS provisioning capabilities of the proposed schedulers.

The rest of the paper is organized as follows. Section II discusses the related research. Section III introduces the system model. Fair STDMA scheduling in directional mmWave WPANs/WLANs is formulated as an optimization problem in Section IV. Section V presents the developed heuristic schedulers. Section VI evaluates the performance of the proposed schedulers. Finally, conclusions are drawn in Section VII.

II. Related Work

STDMA scheduling has been studied in different types of mmWave networks. To maximize the QoS satisfied flows, concurrent transmission scheduling in mmWave WPANs is formulated as an optimization problem in [4]. In addition, to generate a schedule in real-time and satisfy the minimum QoS requirements of flows while minimizing the fractional flows, a heuristic scheduling method, termed STDMA, is also proposed. In [3], concurrent transmission scheduling methods are proposed for transmission rate measurements as well as actual data transmission. The frame-based scheduling protocol proposed in [5], termed FDMAC, exploits spatial reuse to improve network throughput. The scheduling scheme developed in [2], termed D2DMAC, addresses joint scheduling of radio access and backhaul in small cell mmWave networks.

Fair resource allocation has been an active research area and there exists a considerable amount of research on fair scheduling [6]. Most of these methods assign service tags to packets to define their transmission order according to the fairness constraints. The look-ahead fair queueing model proposed in [6] aims to maximize the fair allocation while limiting short-term unfairness among the flows in ad hoc networks. The packet fair schedulers proposed in [7] target both throughput maximization and fairness at the same time.

In contrast to the existing literature on STDMA scheduling, in this paper, we formulate the fair scheduling in mmWave WPANs/WLANs as an optimization problem to maximize the network throughput while ensuring fairness among the flows. In contrast to the existing fair scheduling methods, our heuristic schedulers do not use any tags and generate near-optimal fair schedules in real-time.

III. System Model

The considered network architecture consists of a central coordinator and a number of nodes or wireless devices (WDs), similar to a typical WPAN/WLAN. The central coordinator functions similar to a pico net coordinator (PNC) and coordinates channel access among the WDs so that they can communicate directly in a peer-to-peer fashion. The PNC is deployed with a directional antenna capable of forming β non-overlapping beams, where the beamwidth of each of these beams is $2\pi/\beta$ radians. The other nodes in the network are deployed with steerable directional antennas. We assume an ideal flat-top antenna model for the directional antennas. With this model, the antenna gain is constant within the beamwidth and zero outside the beamwidth [8].

All nodes are half-duplex and adjacent links cannot transmit simultaneously. Each node learns about its neighbors using neighbor discovery techniques [9], and then shares this information with the PNC which in turn uses this information to obtain the network topology. The PNC can obtain the locations of other nodes using wireless channel signatures [10]. It identifies the interfering links of each link using the beamforming information communicated by the WDs. For links (i, j) and (x, y), a variable $h_{xy,ij}$ indicates whether these two links interfere with each other or not. If i and x are outside the beamwidth of y and j, respectively, or i (x) is within the beamwidth of y (j) but does not direct its beam towards y (j), then links (i, j) and (x, y) cannot interfere, and hence, $h_{xy,ij}$ is set to 0; otherwise, $h_{xy,ij}$ is set to 1. The number of interfering links of link (i, j) is called its *interference* [11].

The network time is divided into superframes. As shown in Figure 1, each superframe is subdivided into three periods: beacon, contention access and channel time allocation. Beacon periods (BPs) are for the PNC to transmit scheduling and synchronization information in the network. Each contention access period (CAP) is divided into a number of mini-slots, which are used by the nodes in the network to send transmission requests to the PNC. To accommodate the received transmission requests, the PNC generates a schedule for the upcoming superframe. Following the schedule communicated by the PNC, nodes transmit data during each channel time allocation period (CTAP). During BPs and CAPs, the PNC switches on all its beams and only one beam during CTAPs.

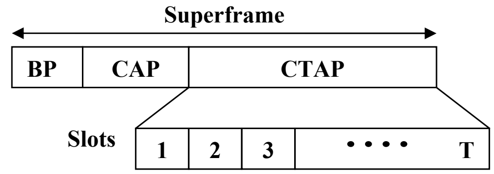


Fig. 1. Frame structure [4].

Due to high path loss in the mmWave band, non-line-of-sight transmissions cannot support high data rates, and hence only LOS transmissions are considered. To compute the transmission rate of a link, we use the interference model given in [4]. When two links (i, j) and (x, y) transmit concurrently, the power received by j from x is computed as: $P_r^{xy,ij} = h_{xy,ij} K P_\tau l_{xj}^{-\alpha}$, where P_τ is the transmission power, l_{xj} is the distance between x and j, $K = 10^{\sigma(l_{ref})/10}$ is the scaling factor corresponding to the reference path loss $\sigma(l_{ref})$ observed at a reference distance l_{ref}, and α is the path loss exponent. Now, the signal to interference plus noise ratio (SINR) of link (i, j) can be computed as:

$$SINR_{ij} = \frac{K P_\tau l_{ij}^{-\alpha}}{N_0 W + \rho \sum_{x \neq i,j} \sum_{y \neq i,j} h_{xy,ij} K P_\tau l_{xj}^{-\alpha}}, \quad (1)$$

where W is the bandwidth, N_0 is the one-side power spectral density of white Gaussian noise, ρ denotes the multi-user interference factor, and (x, y) is a representative of those links that are scheduled for concurrent transmission with (i, j). Link (i, j) can transmit concurrently with the other links provided $SINR_{ij} \geq SINR_{min}(c_{ij})$, where $SINR_{min}(c_{ij})$ is the minimum required SINR to support the transmission rate c_{ij} of link (i, j) [11].

IV. Fair Scheduling in mmWave Networks

The traffic demand of flow f running over the link (i, j) is represented as d_{ij} and c_{ij} represents the transmission rate of link (i, j). Each flow f corresponds to a single-hop link (i, j). The minimum throughput requirement of flow f is called its QoS requirement and denoted as Υ_{ij} $(\leq d_{ij})$. Each flow f is assigned with a weight ω_{ij}, which represents flow f's priority in comparison to the other flows. Flows with the same weight are grouped into one service class and all flows in the network can be categorized into k service classes Ψ_1, Ψ_2, $\cdots$, Ψ_k.

The STDMA schedule of each frame consists of a collection of pairings. In each pairing t, a set of flows, S^t, is scheduled for transmission, and t lasts for a number of slots called the duration, θ_t, of the pairing t. For a network with n nodes, S^t is a $n \times n$ matrix. s_{ij}^t is the element at position (i, j) in S^t. It is set to 1 if link (i, j) is scheduled for transmission in the t-th pairing, otherwise it is set to 0. Hence, the STDMA schedule of a frame can be represented as [11]:

$$\Pi = \theta_1 S^1 + \theta_2 S^2 + \cdots + \theta_m S^m. \quad (2)$$

The optimal fair scheduler should maximize the network throughput and at the same time, should allocate the available resources to various flows in such a way that the service received by each flow is in proportion to its weight. To obtain such a fair schedule, the optimal fair scheduling in mmWave networks can be formulated as a multi-objective optimization problem (**P1**) as follows:

$$\min_{\theta,S}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\lceil d_{ij}/c_{ij}\rceil-\Big(\sum_{t=1}^{m}\theta_t s_{ij}^t\Big)\right),\right.$$

$$\left.\sum_{(i,j)}\sum_{(x,y)\neq(i,j)}\left|\frac{\sum_{t=1}^{m}\theta_t s_{ij}^t c_{ij}}{\omega_{ij}}-\frac{\sum_{t=1}^{m}\theta_t s_{xy}^t c_{xy}}{\omega_{xy}}\right|\right], \quad (3)$$

such that

$$\sum_{t=1}^{m}\theta_t \le T, \quad \sum_{j=1}^{n}(s_{ij}^t+s_{ji}^t)\le 1 \ \forall\ i,t \quad (4)$$

$$\sum_{t=1}^{m}s_{ij}^t \begin{cases}\ge 0,\ \le\lceil d_{ij}/c_{ij}\rceil, & \text{if } d_{ij}>0\\ =0, & \text{if } d_{ij}=0\end{cases}\forall\ i,j \quad (5)$$

$$s_{ij}^t\in\begin{cases}\{0,1\}, & \text{if } d_{ij}>0\\ \{0\}, & \text{if } d_{ij}=0\end{cases}\forall\ i,j,t \quad (6)$$

$$\sum_{t=1}^{m}\theta_t s_{ij}^t\begin{cases}\le\lceil d_{ij}/c_{ij}\rceil, & \text{if } d_{ij}>0\\ =0, & \text{if } d_{ij}=0\end{cases}\forall\ i,j \quad (7)$$

$$\sum_{t=1}^{m}\theta_t s_{ij}^t\begin{cases}\ge\lceil \Upsilon_{ij}/c_{ij}\rceil, & \text{if } d_{ij}>0\\ =0, & \text{if } d_{ij}=0\end{cases}\forall\ i,j \quad (8)$$

$$SINR_{ij}^t \ge SINR_{min}(c_{ij})\ \forall\ i,j,t. \quad (9)$$

Condition (4) indicates that the total duration of the schedule should not exceed the number of slots available in a frame and adjacent links cannot be scheduled in the same pairing. Conditions (5) to (8) indicate that the bandwidth is allocated to only those flows that have pending demands, and the bandwidth allocated to a flow may span over multiple pairings. Condition (8) indicates that each flow has a minimum throughput requirement. Condition (9) indicates that the SINR values of all links scheduled in a pairing should be more than or equal to their minimum required SINR values.

Constraints (7) to (9) and the objective function consist of nonlinear terms and (**P1**) has both integer and binary variables. Thus, problem (**P1**) is a multi-objective mixed integer nonlinear programming problem. It is hard to solve problem (**P1**) in real-time. After linearizing the nonlinear terms in constraints (7) to (9) and the objective function using relaxation techniques [12], if we solve problem (**P1**) using lexicographic methods, it takes a long time to generate a schedule even for a small networking scenario with 5 to 10 flows (results are given Section VI). Hence, to obtain a near optimal schedule in real-time, we develop heuristic techniques.

V. Fair Scheduling Methods

This section presents our proposed fair schedulers and the goal of these schedulers is two fold: (1) first, satisfy the minimum throughput requirements of as many flows as possible; (2) then allocate the remaining bandwidth to various flows while maintaining fairness among them. These methods differ in their fairness provisioning capability. The first scheduler is called "Cumulative Service Based Fair Scheduler (CS-FS)", and it considers the cumulative service received by the flows in their lifetime while performing fair allocation. The second scheduler is called "Partial Service Based Fair Scheduler (PS-FS)", and it aims to maintain fairness among the flows in terms of the service they receive in each individual frame.

We maintain a set of variables for each flow. Δ_f denotes the cumulative service received by flow f in its lifetime, and δ_f denotes the service it has received in the frame under consideration. The generated pairings are maintained in Π. Each entry p in Π has two fields: the list of links scheduled in p is maintained in S_p and θ_p denotes the duration of p.

Before performing scheduling for a frame, the PNC resets the δ values of all flows to 0. If the PNC does not receive a demand from flow f with non-zero cumulative service, then it resets Δ_f to 0. To treat the newly arriving flows similar to the existing flows (that is, the flows with non-zero cumulative services), for each newly arrived flow $g\in\Psi_k$, its cumulative service is set as: $\Delta_g=\frac{\Delta_{f_1}\omega_g}{\omega_{f_1}}$, where $f_1\in\Psi_k$ is the flow that has received the least cumulative service among all flows in class Ψ_k. In case g is the only flow in its service class, then f_1 denotes the flow that has received the least service in the network. F denotes the set of all flows with demands. In the remaining discussion, d_f, ω_f and Υ_f denote the demand (in slots), weight, and QoS requirement (in slots) of flow f, respectively.

A. CS-FS

1) Phase-I: To satisfy the minimum QoS requirements of a large number of flows, the PNC considers flows for scheduling in the increasing order of their interference [11]. The flow $f^{'}$ with the least interference is identified and a pairing $p^{'}$ is generated with $f^{'}$. Then, the remaining flows are considered in the increasing order of their interference and scheduled in $p^{'}$ if their inclusion do not violate the SINR requirements of the flows in $S_{p'}$. After testing all flows, from the flows in $S_{p'}$, the flow with the least QoS requirement (Υ_{min}) is identified and $\theta_{p'}$ is set as: $\theta_{p'}=\min(T,\Upsilon_{min})$, and T is updated as $(T-\theta_{p'})$. Now, the Υ and δ values of each flow $g\in S_{p'}$ are updated as: $\Upsilon_g\ {-}{=}\ \theta_{p'}$ and $\delta_g\ {+}{=}\ \theta_{p'}$. Then $p^{'}$ is pushed onto Π.

The process explained above is repeated until the minimum throughput requirements of all flows are satisfied successfully or there are no more free slots left. Now, for each flow f that is scheduled in this phase, its demand and cumulative service are updated as $(d_f-\delta_f)$ and $(\Delta_f+\delta_f)$, respectively.

2) Phase-II: In this phase, the pairings generated in Phase-I are extended while ensuring fairness of various flows. Flows are considered in the increasing order of their normalized cumulative service values, that is, Δ/ω values, and included in all possible pairings in Π. Whenever a flow $f^{'}$ is scheduled in a pairing $p^{'}$, its pending demand $d_{f'}$ and the cumulative and instantaneous services, that is, $\Delta_{f'}$ and $\delta_{f'}$ are updated as follows. If $d_{f'}\ge\theta_{p'}$, then $d_{f'}$, $\delta_{f'}$ and $\Delta_{f'}$ are updated as $(d_{f'} - \theta_{p'})$, $(\delta_{f'} + \theta_{p'})$ and $(\Delta_{f'} + \theta_{p'})$, respectively. Otherwise, $\delta_{f'}$, $\Delta_{f'}$ and $d_{f'}$ are updated as $(\delta_{f'} + d_{f'})$, $(\Delta_{f'} + d_{f'})$, and 0, respectively.

3) Phase-III: In this phase, CS-FS gives prioritized service to those flows that have received a lower cumulative service in comparison with the other flows. The flow g that has received the highest normalized cumulative service is identified. In comparison to flow g, the extra service *to be received* by each flow $g_1 \in (F - \{g\})$ is denoted as μ_{g_1} and set as: $\mu_{g_1} = \min(d_{g_1}, (\lceil \frac{\Delta_g \omega_{g_1}}{\omega_g} \rceil - \Delta_{g_1}))$. To address the μ values of flows, Algorithm 1 with the complexity $O(T|F|^2)$ is executed with F and Π_1, where Π_1 is the pairings generated in this phase. At the beginning of this phase Π_1 set to ϕ. In line 3 of Algorithm 1, the flows are considered in the increasing order of their normalized cumulative services. Each flow f is included in as many pairings as possible until $\mu_f > 0$. For each pairing p in Π_1, flow f is temporarily included in S_p and checked for the violation of SINR requirements of the flows in S_p in lines 5 and 6. If SINR requirements are not violated for any of the flows in S_p, then f is scheduled in p in line 7. In case θ_p is more than μ_f, then μ_f is set to 0 in line 9, and we provide extra service to flow f on top of its μ_f as follows. If $d_f \geq \theta_p$, then d_f, Δ_f, and δ_f are updated as given in line 11. Otherwise, we replace pairing p with two pairings $p_1^{'}$ and $p_1^{''}$ in lines $13 - 16$. Also, Δ_f, δ_f, and d_f are updated as given in line 15. In case $\theta_p \leq \mu_f$, then d_f, Δ_f, δ_f, and μ_f are updated as given in line 19. After checking all pairings in Π_1, if $\mu_f > 0$ and some free slots are available, then a new pairing is generated with f as given in lines $26 - 36$.

4) Phase-IV: In this phase, CS-FS aims to level the services various flows receive in the current frame. It identifies the flow f_1 with the highest $\delta_{f_1}/\omega_{f_1}$ value. In proportion to δ_{f_1}, the services that the other flows should receive are computed and set as their μ values. Specifically, for each flow $g^{'} \in F$, its $\mu_{g'}$ is set as: $\mu_{g'} = \min(d_{g'}, (\lceil \frac{\delta_{f_1} \omega_{g'}}{\omega_{f_1}} \rceil - \delta_{g'}))$. Then, δ values of all flows are reset to zero, and the process given in Algorithm 1 is executed with F and Π_1. Phase-IV is recursively executed until there is no change in the μ values of all flows.

5) Phase-V: If there still exist pending demands and free slots, then the flow f with the least pending demand is identified and its μ value is set as $\mu_f = d_f$. In proportion to μ_f, the μ values of the other flows are set as: $\mu_{g'} = \min (d_{g'}, \lceil \frac{\mu_f \omega_{g'}}{\omega_f} \rceil) \ \forall \ g^{'} \in (F - \{f\})$, and the process given in Algorithm 1 is executed. Phase-V is recursively executed until all demands are addressed successfully or no more free slots are left. At the end of Phase-V, Π is set as: $\Pi = (\Pi \cup \Pi_1)$.

B. PS-FS

1) Phase-I: Similar to CS-FS, in this phase, PS-FS addresses the minimum throughput requirements of as many flows as possible. First, the flow f_1 with the least interference is identified and a new pairing p_1 is generated with f_1. Then, the flow f_2 with the next least interference is identified. If, after the inclusion of f_2 in S_{p_1}, the SINR values of all flows in S_{p_1} are more than or equal to their minimum required SINR values, then f_2 is included in S_{p_1}; otherwise, not. Similarly the other flows are considered for inclusion in p_1 in the increasing order of their interference. Now, to set the duration of p_1, three cases are possible.

- Among the flows scheduled in p_1, the highest minimum QoS requirement (Υ_{max}) is identified. If $T \geq \Upsilon_{max}$, then θ_{p_1} is set as Υ_{max}, and T is updated as $(T - \Upsilon_{max})$.
- In case Υ_{max} is more than T, then the least minimum throughput requirement (Υ_{min}) among the flows in S_{p_1} is identified. If $T \geq \Upsilon_{min}$, then θ_{p_1} is set as Υ_{min} and T is updated as $(T - \Upsilon_{min})$.
- If $T < \Upsilon_{min}$, then we check whether by preempting some slots from one of the already generated pairings, and allocating these slots to pairing p_1, the number of QoS satisfied flows can be increased or not. For each pairing $p \in \Pi$, the duration is temporarily decreased by $(\Upsilon_{min} - T)$ and the number of flows in p whose minimum QoS requirements are disturbed due to this decrement is obtained. Let $p^{'} \in \Pi$ be the pairing with the least number of disturbed flows, that is, n_1. Assume that, when we set θ_{p_1} to Υ_{min}, the number of QoS satisfied flows in p_1 is n_2. If $n_2 \leq n_1$, then θ_{p_1} is set as 0. Otherwise, the duration of $p^{'}$ is updated as $(\theta_{p'} - (\Upsilon_{min} - T))$, θ_{p_1} is set as Υ_{min}, and T is updated as 0.

If $\theta_{p_1} > 0$, then for each flow f scheduled in p_1, its δ_f, Δ_f, and d_f are updated as $(\delta_f + \min(d_f, \theta_{p_1}))$, $(\Delta_f + \min(d_f, \theta_{p_1}))$, and $\max(0, (d_f - \theta_{p_1}))$, respectively; and then p_1 is included in Π. The process explained above is repeated until the minimum throughput requirements of all flows are satisfied or a pairing with zero duration is generated.

2) Phase-II & III: The working procedures of Phase-II and III of PS-FS are the same as the working procedures of Phase-IV and V of CS-FS, respectively. In Phase-II, PS-FS initializes Π_1 with Π and executes the process explained in Phase-IV of CS-FS. Then, in Phase-III, it executes the process explained in Phase-V of CS-FS. At the end of Phase-III, Π is set as Π_1.

VI. SIMULATION RESULTS

For performance evaluation of the proposed methods, we consider a mmWave WPAN deployed in the 60 GHz band to support high data rate indoor services such as smart classrooms, conference rooms, video streaming, etc. We use the simulation parameters that are most commonly used in the existing literature on mmWave WPANs [4] and are as follows. The PNC is located at the center of a circular region with a radius of 10 m. 80 nodes are distributed at randomly selected locations in the considered region. Each node is deployed with a directional antenna and the antenna beamwidth is set as 60^o. All nodes use the same transmission power, that is, 0.1 mW. The system bandwidth is 1200 MHz. The reference path loss at the reference distance of 1.5 m is 71.5 dB and the path loss exponent is 2 [13]. The background noise level is -134 dBm/MHz [14]. Number of slots in each CTAP is 1000 and the duration of each slot is 18 μs [4]. The presented results are averaged over 50 simulation runs.

Algorithm 1 Stage-III of CS-FS

BEGIN:

initialization: F_1 is the set of flows with non-zero μ values;
while $F_1 \neq \phi$ **do**
  find $f \in F_1$ with the least Δ_f/ω_f; $F_1 = F_1 - \{f\}$;
  for all p in Π_1 **do**
    $S_{p_{temp}} = S_p \cup \{f\}$;
    if $SINR$ values of all flows in $S_{p_{temp}}$ are more than or equal to their $SINR_{min}$ values **then**
      $S_p = S_{p_{temp}}$;
      if $\theta_p > \mu_f$ **then**
        $\mu_f = 0$;
        if $d_f \geq \theta_p$ **then**
          $d_f -= \theta_p$; $\delta_f += \theta_p$; $\Delta_f += \theta_p$;
        else
          Generate two pairings $p_1^{'}$ and $p_1^{''}$;
          $S_{p_1^{'}} = S_p - \{f\}$; $\theta_{p_1^{'}} = \theta_p - d_f$; $S_{p_1^{''}} = S_p$;
          $\theta_{p_1^{''}} = d_f$; $\delta_f += d_f$; $\Delta_f += d_f$; $d_f = 0$;
          remove p from Π_1; push $p_1^{'}$ and $p_1^{''}$ onto Π_1;
        end if
      else
        $d_f -= \theta_p$; $\Delta_f += \theta_p$; $\delta_f += \theta_p$; $\mu_f -= \theta_p$;
      end if
    end if
    if $\mu_f == 0$ **then**
      break;
    end if
  end for
  if $\mu_f > 0$ and $T > 0$ **then**
    Generate a new pairing p_n; $S_{p_n} = \{f\}$;
    if $T \geq \mu_f$ **then**
      $\theta_{p_n} = \mu_f$; $d_f -= \mu_f$; $\Delta_f += \mu_f$; $\delta_f += \mu_f$;
      $T -= \mu_f$; $\mu_f = 0$;
    else
      $\theta_{p_n} = T$; $\mu_f -= T$; $d_f -= T$; $\Delta_f += T$;
      $\delta_f += T$; $T = 0$;
    end if
    push p_n onto to Π_1;
  end if
end while

END;

Flows are established between randomly selected node pairs. The data rate requirements of these flows are uniformly distributed between 1.5 and 3.5 Gbps. Each flow is assigned a weight randomly selected from the group of considered weights: 0.4, 0.3, 0.2, and 0.1. Hence, all flows can be categorized into four service classes.

For performance comparison of CS-FS and PS-FS, the scheduling methods *STDMA* [4] and FDMAC [5] are considered. Network throughput and fairness index are considered as the performance evaluation metrics. The fairness of each service class is evaluated using Jain's fairness index [15], defined as: $Fairness(\Psi) = (\sum_{f\in\Psi} a_f)^2/(N_\Psi \sum_{f\in\Psi}(a_f)^2)$, where a_f is the total number of slots allocated to flow f in a simulation run, and N_Ψ is the number of flows in class Ψ.

Figures 2(a) to 2(d) show the fairness indices of the four service classes as the number flows vary from 15 to 50. When the number of demanding flows is around 20, the difference in the performance of the proposed methods and STDMA and FDMAC is not very different, since all four methods accommodate almost the same traffic. However, as the number of demanding flows increases, the addressed traffic with both STDMA and FDMAC does not increase much. Also, there may exists a huge difference in the service received by any two flows from the same service class. On the other hand, the addressed traffic with the proposed methods increases until the number of flows reaches up to 50. In addition, irrespective of the number of demanding flows, CS-FS and PS-FS allocate the available limited resources to various flows such that the fairness of each service class improves. Thus, their fairness indices are much better, compared to STDMA and FDMAC.

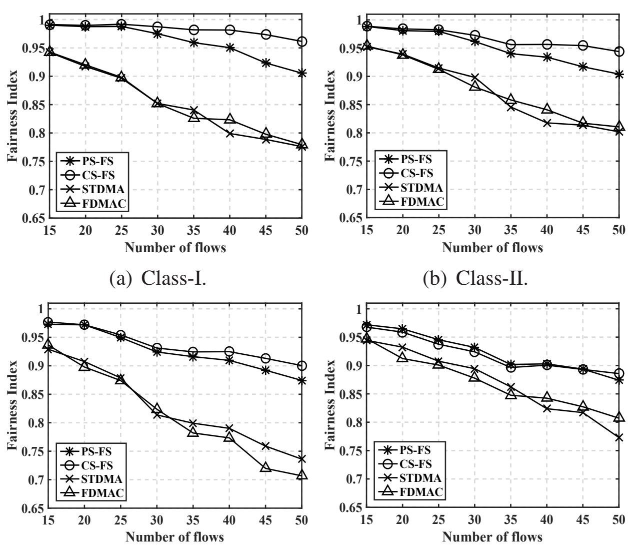
Figure 3 shows the network throughput of all four methods as the number of flows vary from 15 to 50. For STDMA, the set of transmitting flows can only change when the addition of a new flow into the set of scheduled flows contributes towards the throughput enhancement. Hence, the spatial reuse may be under utilized when highly contending flows are scheduled together. For FDMAC, the flows scheduled in a pairing remain active until the demands of all scheduled flows are addressed completely, and thus, in some slots the spatial reuse may be exploited partially. On the other hand, in their first phase, both CS-FS and PS-FS use interference information of flows to maximize the number of QoS-satisfied flows. Then, in the following phases, more flows are scheduled in the pairings generated in Phase-I. As a result, the proposed methods achieve a higher throughput, compared to STDMA and FDMAC.

As the number of demanding flows increases, the variance in the received services of flows increases owing to the increased variance in flow interference. Thus, we observe a slightly decreasing trend in the fairness indices with all methods. However, the proposed methods, by considering both long-term and short-term resource utilization of flows during scheduling, achieve much better fairness indices, compared to STDMA and FDMAC. In its first phase, PS-FS may set the duration of some of the pairings as the maximum QoS requirement among the flows scheduled in these pairings. Hence, sometimes, PS-FS does not exploit the spatial reuse completely, which also leads to unfair utilization of resources by the flows. As a result, its fairness indices are lower, compared to CS-FS.

The performance of CS-FS is also compared with the optimal solution obtained using an open-source tool YALMIP [16] and a trial version of the MOSEK solver [17]. Since it takes a long time to obtain the optimal solution, a smaller networking scenario where the number of flows increases from 5 to 10 is considered for performance evaluation. The weight of each flow is either 10.0, 5.0 or 3.4, and the results are averaged over 5 simulation runs.

To obtain the optimal solution, first the nonlinear terms of problem (**P1**) are linearized using relaxation techniques [12], and then, (**P1**) is solved using lexicographic method. With this method, first problem (**P1**) is solved for the first objective to obtained an optimal value for the pending demand (pd) at the end of scheduling. Next, a constraint that bounds the value of first objective with pd is added to problem (**P1**), and then, (**P1**) is solved for the second objective.

Figure 4 shows a comparison of the throughput of CS-FS and the optimal solution and Figure 5 shows the fairness indices of the class with weight 10.0. CS-FS obtains almost similar throughput as that of the optimal solution. In the case of fairness sometimes it provides a better fairness due to the following fact. In Phase-I, CS-FS maximizes the number QoS satisfied flows, and consequently, the network throughput. But, from Phase-II onwards, its priority is fair allocation rather than network throughput. On the other hand, the optimal solution is obtained while constraining the throughput to a maximal value. Thus, the fairness of the optimal solution is sometimes slightly lower than that of CS-FS. The execution time of the optimal solution is much higher, compared to CS-FS, and increases with the number of flows, as shown in Figure 6. These results show that CS-FS obtains near optimal solution in real-time.



(a) Class-I. (b) Class-II.

(c) Class-III. (d) Class-IV.

Fig. 2. Number of flows vs fairness index.

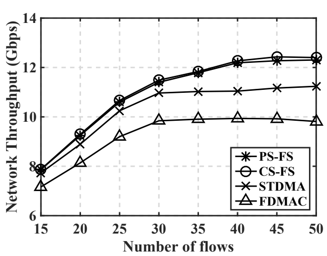


Fig. 3. Number of flows vs network throughput.

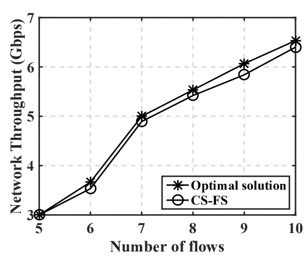


Fig. 4. Number of flows vs network throughput.

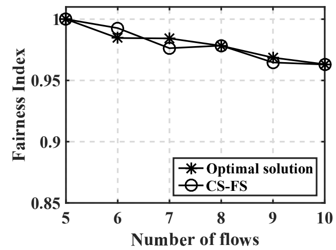


Fig. 5. Number of flows vs fairness index of Class-I.

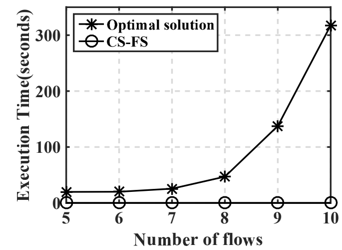


Fig. 6. Number of flows vs execution time.

VII. Conclusions

This paper addresses fair concurrent transmission scheduling in mmWave WPANs/WLANs. To provide prioritized service to flows with different weights, STDMA scheduling in mmWave networks is formulated as a multi-objective optimization problem. The developed heuristic STDMA schedulers first address the QoS requirements of all possible flows, and then allocate the remaining bandwidth while considering the short-term and long-term service received by the flows. Simulation results from a mmWave WPAN show that the proposed methods can achieve a higher throughput while providing a better fairness, compared to the existing STDMA scheduling methods.

References

[1] "Qualcomm data challenge", https://www.qualcomm.com/documents/rising-meet-1000x-mobile-data-challenge.

[2] Y. Niu, C. Gao, Y. Li, L. Su, D. Jin and A. V. Vasilakos, "Exploiting device-to-device communications in joint scheduling of access and backhaul for mmWave small cells", *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2052-2069, 2015.

[3] Y. Niu, Y. Li, D. Jin, L. Su and D. O. Wu, "A two stage approach for channel transmission rate aware scheduling in directional mmWave WPANs", *Wirel. Commun. Mob. Com.*, vol. 16, no. 3, pp. 313-329, 2016.

[4] J. Qiao, L. X. Cai, X. Shen and J. W. Mark, "STDMA-based scheduling algorithm for concurrent transmissions in directional millimeter wave networks", *Proc. IEEE ICC*, pp. 5221-5225, 2012.

[5] I. K. Son, S. Mao, M. X. Gong and Y. Li, "On frame-based scheduling for directional mmWave WPANs", *Proc. IEEE INFOCOM*, pp. 2149-2157, 2012.

[6] H. Luo and S. Lu, "A topology-independent fair queueing model in ad hoc wireless networks", *Proc. IEEE ICNP*, pp. 325-335, 2000.

[7] H. Luo, S. Lu, and V. Bharghavan, "A new model for packet scheduling in multihop wireless networks", *Proc. ACM MOBICOM'00*, pp. 76-86, 2000.

[8] R. Mudumbai, S. K. Singh and U. Madhow, "Medium access control for 60 GHz outdoor mesh networks with highly directional links" *Proc. IEEE INFOCOM*, pp. 2871-2875, 2009.

[9] F. Yildirim and H. Liu, "A cross-layer neighbor-discovery algorithm for directional 60-GHz networks", *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4598-4604, 2009.

[10] T. S. Rappaport, J. H. Reed and B. D. Woerner, "Position location using wireless communications on highways of the future", *IEEE Communications Magazine*, vol. 34, no. 10, pp. 33-41, 1996.

[11] L. R. Lakshmi and B. Sikdar, "STDMA scheduling for WLANs and WPANs with non-uniform traffic demand", *Proc. IEEE LANMAN'17*, pp. 1-6, 2017.

[12] H. D. Sherali and W. P. Adams, "A reformulation-linearization technique for solving discrete and continuous nonconvex problems", Kluwer Academic: Boston, MA, 1999.

[13] N. Moraitis and P. Constantinou, "Indoor channel measurements and characterization at 60 GHz for wireless local area network applications", *IEEE Trans. Antennas Propag.*, vol. 52, no. 12, pp. 3180-3189, 2004.

[14] IEEE P802.15-05-0493-27-003c, "TG3c Selection Criteria", Jan., 2007.

[15] R. Jain, A. Durresi, and G. Babic, "Throughput fairness index: an explanation", *ATM Forum Document Number: ATM Forum/990045*, Feb. 1999.

[16] YALMIP Wiki [Online]. Available: http://users.isy.liu.se/johan1/yalmip/.

[17] MOSEK Wiki [Online]. Available: https://www.mosek.com.