An Efficient Scheduling Scheme For MIMO Wireless Mesh Networks With Fairness Constraints

Abdelhalim Driouech, Wessam Ajib, Elmahdi Driouech
Université du Québec à Montréal
{driouech.abdelhalim,driouech.el mahdi}@courrier.uqam.ca, ajib.wessam@uqam.ca

Abstract

Multi hop wireless mesh networks presents a promising solution to extend the coverage and increase the number of clients sharing the same broadband connection. Introducing the MIMO technology at the physical layer of those networks allows to improve the performance and support a higher number of clients. However, in such systems the absence of an efficient and fair medium access strategy can lead to severe unfairness between the clients and to poor system performances. In this paper, we propose a fair scheduling algorithm that improves considerably the bandwidth utilization while the fairness is guaranteed. We evaluate our proposed algorithm by simulation in order to show the maximum throughput that can be obtained while the different nodes of the MIMO-based wireless mesh networks are served fairly.

1 Introduction

The increasing demand for high coverage of wireless local area networks (WLAN) is driving the installation of very large number of access points. In the infrastructure-based wireless networks, the access points use wired connections, which introduce high cost and complexity. Multi-hop wireless mesh networks (WMN) allows multi-hop radio transmissions between the access points, which permits to extend the coverage and to simplify the deployment. Hence, this type of networks becomes unavoidable for the next-generation of wireless networks. In a WMN there is at least one hot spot (HS) connected to the internet and many transient access points (TAPs) that use the wireless medium to transfer their clients traffic to and from the “wired” hot spot. The WMN can relay, at the need, the traffic of large categories of consumers from the simple residences Internet connection to industrial and commercial venues. In [1], Akyildiz et al. present a survey on advances and open research issues in WMNs.

The need to increase the network capacity becomes critical for satisfying the increasing number of users and supporting the new large-bandwidth applications. Hence, we need to optimize the utilization of the bandwidth which is a limited resource in wireless systems. The multiple-input multiple-output (MIMO) wireless systems are emerging as an interesting solution for improving the bandwidth utilization. In [2], the authors show the large benefits that can be obtained from the diversity and the spatial multiplexing gains of MIMO systems. Hence, MIMO technology is an important candidate to solve the problem of bottleneck in wireless backhaul networks when it is used with an efficient access medium mechanism.

We define the throughput that can be achieved by a WMN network as the amount of data that can be transported by the TAPs from the clients to the HS or inversely form the HS to the clients. To provide high throughput, MAC protocols have to be adapted to the particularities of the multi-hop mesh networks. In [3], the authors investigate the effect of the distribution of antennas on the capacity in a multi hop wireless mesh network. They show that the bottleneck depends on the transmission configuration (the number of antennas used on each node). In [4], the impact of simultaneous transmissions in WMN network is studied and a new MAC protocol that maximizing the network throughput is investigated. Unfortunately, most of the existing MAC protocols aims only to maximize the throughput without taking into consideration the fairness between the different access points. Furthermore, the users located more than one hop away from the HS may suffer from starvation. Hence the maximization of the WMN throughput should be constrained by the fairness.

In [5], the impact of providing fairness on the system capacity is investigated. In [6], the authors proposes an algorithm for enhancing fairness constrained system capacity. The differences between our work and the previously mentioned ones is that we adopt the utilization of multiple antennas and the use of spatial reuse between all the links (In [6], the authors consider spatial reuse only for the non contention links).
In this paper, we propose a scheduling algorithm to be used in MIMO wireless mesh networks with different topologies. We consider two kinds of topologies: chain and grid. Anyhow, our proposition can be applied to other network topologies. The proposed algorithm focuses on increasing the total network throughput in a WMN system constrained by per-node fairness (the amount of traffic generated at each node is assumed equal). In [7], it is proven that the link-based scheduling is more convenient for networks with a high density of traffic, which is our case (maximizing the network throughput). To attain the scheduler objectives, first, our algorithm maximizes the per link throughput without taking into account the per node fairness. As a second step, our proposed algorithm introduces the fairness objective to ensure that all the network links almost achieve the same throughput.

This paper is organized as follows. We present in Section 2 the system model including the network, access, interference and channel models. In Section 3, we introduce the scheduling problem formulation, and detail our scheduling algorithm. Numerical results, provided in section 4, evaluate the performance of the proposed scheduler in terms of both fairness and capacity enhancement and show the benefits of the new proposed scheduling scheme. Finally, we conclude our paper in Section 5.

2 System Model

2.1 Network Model

We consider a network of $N$ router nodes ($N - 1$ TAPs and the $N$th node is the HS). The nodes are disposed in two different topologies: chain topology and grid topology. The distances between each two adjacent nodes have the same value which is assumed to be equal to unity. The link between node $n$ ($n = 1, \ldots, N - 1$) and the following node is also indexed “$n$”. Each node has to forward both data coming from the precedent nodes and the traffic generated by its own clients. Without loss of generality, we consider only the traffic from the clients to the HS. For clarity purposes, we assume that the clients are using a different RF channel from the one used by the router nodes. The data received by each node from its clients is forwarded in a multi-hop manner in order to reach the last node $N$. We also assume that each group of clients, belonging to the same node $n$, can generate $G_n$ bits per second. Each node is equipped with $M$ antennas that can be used for transmission or reception. Fig. 1 illustrates a network with 5 nodes disposed in a chain topology. The considered network is assumed to be composed by only one chain.

Fig. 2 illustrates a 9-nodes network using a grid topology.

2.2 Access Model

To access the channel, we utilize time division multiple access (TDMA). Note that other orthogonal multiple access schemes can be used. We also consider the spatial reuse which improves the bandwidth utilization and consequently the system performance. The combination of TDMA and spatial reuse is commonly called Space time division multiple access (STDMA). Time is divided into frames. Each frame contains a different number of time slots (TSs) which are needed to activate all the network links. Note that each link $n$ must be activated exactly during $L_n$ TSs within a single frame where $L_n$ is the number of nodes forwarding their data to node $n$. In the network given by Fig. 2, node 8 has to forward data coming from nodes 4, 5 and 7. Therefore, link 8 must be activated during 4 TSs. The spatial reuse permits the activation of one or more links at the same slot.

2.3 Interference Model

Let $A$ be the set of indices of the activated links during a given TS. We consider all the other transmissions as interference. Hence, each receiver node $n$ experiences a loss of throughput caused by $|A| - 1$ (where $|.|$ denotes the cardinality of a set) interfered signals belonging to the set $I = A\setminus\{n\}$ from other transmitting nodes. Note that this assumption is more realistic than the one considering an interference range equal to once or twice the transmission range that was used in most similar works [3] [6].
2.4 Channel Model

We denote by $H_{n_1 n_2}$ the $M \times M$ channel matrix between nodes $n_1$ and $n_2$ ($n_1 = 1, \ldots, N - 1$, $n_2 = n_1 + 1, \ldots, N$). The channel is given by $H_{n_1 n_2} = \sqrt{\alpha} H_{n_1 n_2}$, where $H_{n_1 n_2}$ is a $M \times M$ matrix whose elements are assumed to be independent identically distributed (i.i.d.) complex Gaussian variables with unit variance and zero mean, $d$ is the distance between nodes $n_1$ and $n_2$, and $\beta$ is the path loss coefficient. We assume that the channels state information (CSI) are fixed during one TS. For each link $n_i$, $H_{n_i n_r}$ is perfectly known at the sender node $n_i$ and the receiver node $n_r$. We also assume that the power available for transmission at each node is fixed to $P$. The received signal at node $n_r$ is given by the $M \times 1$ vector:

$$y_{n_r} = \sqrt{d} H_{n_i n_r} s_n + \sum_{i \in I} \sqrt{d} H_{i,n_r} s_i + w_{n_r}$$

(1)

where $s_n$ is the $M \times 1$ desired signal vector transmitted from node $n$, $s_i$ is the $M \times 1$ interfered signal vector transmitted from node $i$, and $w_{n_r}$ is the $M \times 1$ vector of i.i.d. additive zero mean white complex Gaussian noise with unit variance.

Since the channel matrix is perfectly known at the transmitter, we make use of singular values decomposition (SVD) of the channel matrix. $H_{n_i n_r}$ between nodes $n_i$ and $n_r$ is decomposed as follows:

$$H_{n_i n_r} = U_{n_i} \Sigma \sigma_n V_{n_r}^H$$

(2)

where $U_{n_i}$ and $V_{n_r}$ are $M \times M$ unitary matrices (the $(\cdot)^H$ denotes the conjugate transpose operator) and $\Sigma_n$ is a $M \times M$ diagonal matrix whose diagonal elements are the singular values of the matrix $H_{n_i n_r}$.

Hence, equation (1) can be rewritten as:

$$\tilde{y}_{n_r} = \sqrt{d} \Sigma \sigma_n \tilde{s}_n + \sum_{i \in I} \sqrt{d} U_{n_i}^H H_{i,n_r} V_{n_r}^H \tilde{s}_i + \tilde{w}_{n_r}$$

(3)

where $\tilde{y}_{n_r} = U_{n_r}^H y_{n_r}$, $\tilde{s}_n = V_{n_r}^H s_n$, $\tilde{s}_i = V_{n_r}^H s_i$ and $\tilde{w}_{n_r} = V_{n_r}^H w_{n_r}$.

The SVD decomposes the original channel into a set of parallel Gaussian channels with independent noise terms having identical variances. To optimize the power allocation, we make use of waterfilling to distribute efficiently the available power among the channels. Therefore, the maximum achievable capacity that can be supported by the link $n_r$ in the time slot $t$ is given by:

$$C(n, t) = \sum_{i=1}^{\nu_n} \log_2 \left( 1 + \frac{\sigma_n^{(i)}}{\sigma_n^{(i)} + 1} \right)$$

(4)

where $\nu_n$ is the rank of the matrix $H_{n_i n_r}$, $\sigma_n^{(i)}$ is the $i$th element of the matrix $\Sigma_n$, $P_n^{(i)}$ is the power allocated to the $i$th channel on link $n$ and $M_{i,n_r} = U_{n_i}^H H_{i,n_r} V_{n_r}^H$.

3 Scheduling Algorithm

3.1 Scheduling Problem Formulation

We propose a scheduling algorithm that (i) maximizes the spatial reuse in order to optimize the utilization of the bandwidth, (ii) ensures a per-node fairness, and (iii) enhances the network throughput subject to constraint (ii).

As mentioned earlier, each group of clients can generate and send at most $G_n$ bits per second. Therefore, the per-node fairness is respected if we have:

$$G_n = C_m \quad \forall n, m \in \{1 \ldots N\}$$

(5)

where $G_n = \frac{C_n}{T}$, $T$ is the number of TSs in one frame and $C_n = \sum_{t=1}^{T} C_{n,t}$ is the average rate that can be achieved in one TS by the link $n$ where $L_n$ is the number of nodes forwarding their data to node $n$. Note that if link $n$ is not activated at time slot $t$ then we denote $C(n, t) = 0$.

In this paper, we define the WMN network throughput as the amount of data that can be sent from the clients to the HS. Hence the network throughput (assuming a perfect per-node fairness) can be computed as:

$$R = \sum_{n=1}^{N-1} G_n = \frac{1}{T} \sum_{n=1}^{N-1} C_n$$

(6)

To obtain the highest possible throughput as defined by (6) (assuming that the condition (5) is respected), we have to find a tradeoff between the two following sub-objectives:

1. maximize each $C_n$ separately $\forall n \in \{1 \ldots N - 1\}$;
2. minimize the duration of a frame $T$;

If we average over all the realizations of the channels, it is widely accepted that the spatial reuse increases the capacity. Anyhow, this statement is not true for any realization of the channels. The rational behind our proposition is to select at each realization of the channels the links to be activated (with or without spatial reuse). To ensure per-node fairness, we need to respect a third sub-objective:

3. ensure that $C_n = C_m$ $\forall n, m \in \{1 \ldots N\}$.

Our scheduling problem is finally reduced to the one of satisfying the three above-mentioned sub-objectives. We introduce in the next subsection our proposed scheduling scheme as a solution of the formulated problem.
3.2 The proposed scheduling

To simplify the illustration of our scheduling, we consider a chain topology network. The same algorithm steps can be applied to the grid topology. We start by enumerating all the combinations corresponding to the links activation (denoted configurations). Let \( N_c \) be the number of all possible configurations. For the architecture in Fig. 1 \((N = 5)\), we have the seven configurations \((N_c = 7)\) given in Fig. 3:

![Diagrams showing possible configurations for a chain wireless MIMO mesh network with \(N = 5\)]

**Figure 3. Possible configurations for a chain wireless MIMO mesh network with \(N = 5\)**

The scheduler decision consists in choosing a configuration at each TS and activating the links belonging to the chosen configuration. Remember that each link \( n \) needs to be activated \( n \) times within a frame.

It is obvious that \( C_n \) for any link \( n \) when activated in a without-spatial-reuse (wSR) configuration (Fig. 3.a) is higher than the one given by any spatial-reuse (SR) configuration (Fig. 3.b). The loss of capacity is due to the interference between the simultaneously activated links. On the other hand, to decrease \( T \), the scheduler may choose configurations with more than one activated link. However, the scheduler should choose at each TS only one configuration and this choice should maximize the average rate for each link during one frame. Hence, the scheduler has to know at each TS if it is better to select a wSR or a SR configuration.

We define \( T_{\text{max}} \) as the number of time slots needed in one frame using only wSR configurations. We introduce for each link \( n \) a set \( V_n \) containing the indices of the SR configurations that activate the link \( n \) (e.g. in Fig. 3, \( V_1 = \{5, 6\} \) for link \( n = 1 \)). We also define for each link \( n \) a \( N_c \times 1 \) vector denoted by \( z_n \). Each element of this vector \( z_n(n_c) \), \( n_c = 1, \ldots, N_c \) represents the maximum throughput that can be achieved by link \( n \) using the configuration \( n_c \). The value of \( z_n(n_c) \) is computed by (4) if the link \( n \) is activated in the configuration \( n_c \) (i.e., \( n_c \in V_n \)) and it is equal to zero otherwise.

Note that for a chain topology, each link \( n \) is activated exactly during \( n \) TSs within one frame, i.e. \( L_n = n \). We introduce for each link \( n \) a \((N_c - N + 1) \times 1 \) vector denoted by \( g_n \). Each element of this vector \( g_n(n_c) \), \( n_c = N, \ldots, N_c \) represents the gain of the per slot capacity when choosing the SR configuration \( n_c \) compared to the wSR configuration that activate the same link \( n \), and it is computed as

\[
g_n(n_c) = \begin{cases} g_n(n_c) & \text{if } n_c \in V_n \\ 0 & \text{otherwise} \end{cases} \tag{7}\]

where \( a \) is the \((N_c - N + 1) \times 1 \) vector that contains the number of links that would be activated in each SR configuration. Note that \( a(n_c) = 1 \) represent the gain in term of TSs between configurations \( n \) and \( n_c \) (e.g., in Fig. 3, \( a(n_c = 6) = 2 \)).

The next step is to define a \((N - 1) \times 1 \) vector (denoted \( b(n) \)) where the element \( b(n) \) \((n = 1, \ldots, N - 1)\) is the index of the configuration that gives the best gain in term of capacity for link \( n \). The index of this configuration is computed as follows:

\[
b(n) = \begin{cases} n & \text{if } \min_{n_c} g_n(n_c) = \max_{n_c} |g_n(n_c)| \\ \arg \max_{n_c} |g_n(n_c)| & \text{otherwise} \end{cases} \tag{8}\]

Note that \( n_c \) must be in \( V_n \).

Since we can activate only one configuration at each TS, the index of the chosen configuration is the element \( b(n) \) having the biggest gain \( g_n(b(n)) \). This configuration allows us to ensure the best throughput for the link \( n \). However, if the selected configuration activates more than one link then the other links may experience a low throughput. Therefore, in the worst case, those other links may be at a disadvantage (e.g. if at time slot \( t \), the configuration 7 is the best for link 1, then the link 4 can be at a disadvantage if at \( t \) its achievable throughput at configuration 7 is low). Hence, before choosing a configuration, we should ensure that it will not disadvantage any of the activated links. In the following, we will modify our scheduler to eliminate this problem by considering the third sub-objective.

In order to guarantee a per-node fairness, we have to modify our configuration selection to avoid disadvantaging any link. Therefore, we consider only the configurations \( b(n) \) where link \( n \) achieves the lowest throughput compared
to the other activated links belonging to the same configuration. The modification will consist in introducing a new \((N - 1) \times 1\) vector, \(\hat{b}\). The new vector contains only the new considered configurations. This vector \(\hat{b}\) is given by:

\[
\hat{b}(n) = \left\{
\begin{array}{ll}
   b(n) & \text{if } z_n(b(n)) = \min_{n' \in \{1, \ldots, N-1\}} z_{n'}(b(n)) \\
   0 & \text{otherwise}
\end{array}
\right.
\]  \hspace{1cm} (9)

Finally, the scheduling algorithm chooses the configuration corresponding to the element \(\hat{b}(n)\) having the biggest gain \(g_n(\hat{b}(n))\). Therefore, this chosen configuration provides the best throughput for link \(n\) while the other activated links are not disadvantaged. It is important to highlight that a configuration that activates an already satisfied link should not be taken into account for the following TSs since each link \(n\) should be activated exactly during \(n\) TSs.

4 Numerical results

![Figure 4. The per-node throughput vs. SNR for a chain topology](image4)

In this section, we analyze the efficiency of the proposed scheduling algorithm in terms of the achievable network throughput through simulations. We also compare the throughput of each node to show the fairness achieved by our scheme. We assume that the path loss coefficient \(\beta\) is equal to 3.5.

In Fig. 4, we show the per-link fairness provided by our scheme for chain topology. We plot the per-node throughput vs. signal to noise ratio (SNR) for a network with \(N = 5\) (the upper curves) and for \(N = 7\) (the lower curves with \(M = 4\) or \(M = 2\)). It can be seen that the throughput is almost the same for all the nodes, which proves the high degree of fairness between the nodes. The standard deviation between the throughputs of the different links is less than 0.02 bps/Hz. Also, when \(N\) increases, the per-node throughput decreases. This is due to the increase in the load of the network since the nodes belonging to bigger networks have to forward more traffic. Fig. 4 shows also the impact of the number of antennas \(M\) on the system performance. In fact, the per-node throughput scales almost linearly with \(M\). This result shows that the proposed scheduling algorithm extracts the advantages of using multiple antennas even for an interference constrained environment. Hence, we can increase the system throughput only by deploying multiple antennas at the nodes, without having to pay any extra bandwidth or power.

Fig. 5 compares, for chain and grid topologies, the proposed algorithm and an opportunistic scheduler using TDMA without spatial reuse. The second scheduler chooses the configuration that activates the link having the highest throughput. Each node is equipped with four antennas \((M = 4)\). We notice that our scheduler outperforms the one without spatial reuse, especially for low SNR. However, in the case of high SNR \((SNR \geq 10\ dB)\), the gap between the performances of the two schemes becomes smaller and may disappear in grid topology. This result is caused by the high impact of interference for high SNR systems which forces our algorithm to choose the wSR configurations. Therefore, the two schemes become almost identical. Also, in the case of a chain topology and for a fixed SNR, the gap between the two compared schemes increases as the number of the network nodes is increased since the
impact of the interference becomes less significant.

![Figure 6. Number of TSs in 1 frame vs. the number of nodes for SNR = 10](image)

Fig. 6 compares between our scheme and the TDMA scheduling without SR in terms of number of time slots needed in one frame. We can clearly see that our scheme outperforms the one without spatial reuse. Indeed, the gap between the duration of a frame for the two compared algorithms becomes larger as $N$ increases.

In Fig. 7, a comparison between our proposed scheme and the optimal one for a chain topology is provided. The optimal scheme performs an exhaustive search over all the possible combinations of configurations assuming that the CSI of all the time slots within one frame is known. We clearly see that our scheme achieves near optimal performance. Also, the gap between the two compared schedulers reduces as SNR increases since the interferences becomes more important and the two schedulers tend to select the wSR configurations.

5 Conclusion

This paper proposed a scheduling algorithm for the multi hop MIMO based wireless mesh networks using the STDMA access model and a per-link based assignment model. Note that our solution focuses on giving an idea about the achievable network throughput. More studies on the complexity and the scalability of the algorithm are needed to have a practical solution. We have shown through simulations that our scheduling enhances the network throughput by deciding when it is beneficial to take advantage from the spatial reuse. The scheduling achieves a per-node fairness by ensuring that no one of the simultaneously activated links will be disadvantaged.

References


