Centralized Scheduling with Sum-Rate Optimization in Flexible Half-Duplex Networks

Shalanika Dayarathna*, Mohammadkhani Razlighi†, Rajitha Senanayake*, Nikola Zlatanov† and Jamie Evans*

*Department of Electrical and Electronic Engineering, University of Melbourne, Australia
†Department of Electrical and Computer Systems Engineering, Monash University, Australia

Abstract—In this paper, we focus on maximization of the instantaneous sum-rate in flexible half-duplex networks, where nodes have the flexibility to choose to either transmit, receive or be silent in a given time slot. Since the corresponding optimization problem is NP-hard, we design low-cost algorithms that give sub-optimal solutions with good performance. We first consider two existing approximation techniques to simplify the sum-rate optimization problem: arithmetic-geometric means inequality and another utilising the tight lower bound approximation. We then propose a novel pattern search algorithm that performs close to exhaustive search but with significantly lower complexity. Comparing the performance of the proposed algorithm with respect to existing resource allocation techniques, we observe that our proposed algorithm provides significant sum-rate gains.

I. INTRODUCTION

In wireless networks, we commonly use two communication modes namely, full-duplex and half-duplex. In half-duplex mode, nodes cannot transmit and receive using the same time/frequency resources. Therefore, transmission and reception must be separated in the time domain (TDD) or in the frequency domain (FDD). In TDD half-duplex networks, a node uses a single frequency band but can only transmit or receive data in a given time slot. Therefore, it is generally pre-defined which time slots are used for transmission and which are used for reception. When the allocation of time slots is not pre-defined then the nodes have the flexibility to select when to transmit and when to receive. This concept is known as flexible duplex. Much research has been conducted to dynamically adjust the uplink/downlink time allocation based on traffic data and channel coefficients. In fact, flexible duplex is one of the key technologies in fifth generation (5G) cellular networks to optimize resource utilization [1]. For example, mode selection and interference aware resource allocation are two important concepts used in device-to-device (D2D) communication underlying cellular communication. A combination of mode selection and interference aware resource allocation can be achieved by re-modelling a D2D network as a flexible duplex network with multiple nodes where each pair of nodes has the choice of selecting the transmission direction in a given time slot. One of the first results in dynamic uplink/downlink resource allocation for TDD half-duplex networks was presented in [2], where the authors adjust the uplink/downlink ratio based on the capacity of a frame. The concept of dynamically adjusting the time allocation for uplink and downlink was introduced in [3] for a network of one base station and a D2D user pair. This work was then extended to a more general multi-link scenario in [4] by adjusting the time allocation to reduce the energy consumption of the network. Papers [5] and [6] use game theory to dynamically adjust uplink/downlink allocation to either minimize overall delay or to maximize overall transmission throughput. Recently, in [7], the authors use a rate splitting strategy to reduce specified uplink/downlink time allocations by introducing a common time allocation that can be used by both uplink and downlink. In all of the above literature, the fixed nature of uplink/downlink allocation is reduced by dynamically adjusting it for the entire transmission period or a time frame. However, it was still fixed for a given time slot.

Uplink and downlink resource allocation for a flexible duplex network is proposed in [8] where the author proposes a successive approximation of fixed point and resource muting for the dominant interferer to reduce the probability of inter-mode interference. The concept in [8] is similar to that used here, however, the problem formulation and optimization algorithm are different. In addition, in this work we have not limited the flexible duplex to be only between cells. Optimization of two-way scheduling with topology graph and dynamic programming is proposed in [9], where the authors consider $M$ pairs of nodes that communicate with each other. This can be considered as a special case of the general model considered in this paper. In [10], the authors present a fast instantaneous signal-to-interference-plus-noise-ratio (SINR) based mode selection for D2D devices within a flexible uplink/downlink TDD cellular network which considered a different approach to solving the optimization problem. The authors of [11] consider a similar approach of developing the system model and instantaneous SINR expression. However, they have limited their optimization problem to maximizing the sum of the SINRs in the network.

In this paper, we consider a flexible half-duplex general network and try to decide which links should be active and in which direction the transmission should take place in an active link in order to maximize the sum-rate of the entire network. Since, we assume channel state information (CSI) is available at a central node, this can be considered as a centralized scheduling technique for flexible half-duplex networks with sum-rate performance optimization. Even though assuming

This work was supported by the Australian Research Council Discovery Project under Grant DP180101205 and Discovery Early Career Researcher Award under Grant DE180100501.
that full CSI is available at a central node is not practical, it provides an upper bound for any distributed scheduling scheme and thus, it is a commonly used assumption in the literature [9]–[11].

II. SYSTEM MODEL

We consider a general flexible half-duplex network consisting of \( K \) nodes that are randomly placed in a network. We assume that each node in the network is impaired by unit-variance additive white Gaussian noise (AWGN), and that the links between the nodes undergo block fading.

Let the transmission on the network be carried-out over \( T \rightarrow \infty \) time slots, where a time slot is small enough such that the fading on all network links can be considered constant during a time slot. Thus, for a given time slot, a signal transmitted from node \( j \) and received at node \( k \) will have a constant normalized magnitude-squared fading gain \( \gamma_{j,k} \).

In addition, we assume that the established connections do not change and new connections do not establish during the transmission period \( T \). As we only consider the instantaneous communication between nodes in this paper, we can define the weighted connectivity matrix for a given time slot as,

\[
G_{j,k} = \begin{cases} 
\gamma_{j,k} & \text{if signal transmitted from node } j \text{ can reach node } k \\
0 & \text{otherwise.}
\end{cases}
\]

We assume that global CSI is available at a central node, where matrix \( G \) is constructed at the start of each time slot.

For a non-dynamic network, desired and interference links do not change during the period. As such, we define matrices for desired and interference links at the beginning of the transmission period as,

\[
W = \begin{cases} 
1 & \text{if node } k \text{ considers transmitted signal from node } j \text{ as desired} \\
0 & \text{otherwise,}
\end{cases}
\]

\[
\bar{W} = \begin{cases} 
1 & \text{if node } k \text{ considers transmitted signal from node } j \text{ as interference} \\
0 & \text{otherwise.}
\end{cases}
\]

In this paper, we have assumed that each node has only one desired receiving signal. Therefore, each column of matrix \( W \) can have only one non-zero entry. Also note that the diagonal entries of \( \bar{W} \) have to be 0 as a given node cannot create a desired link to itself. Then, we obtain desired and interference fading gain matrices, denoted by \( D \) and \( I \), respectively, as

\[
D = G \circ W, \\
I = G \circ \bar{W}.
\]

Let \( d_k \) and \( i_k \) denote the \( k \)-th column vector of matrices \( D \) and \( I \), respectively.

Based on these notations, we present the optimization problem formulation for a general flexible half-duplex network.

III. PROBLEM FORMULATION

In a flexible half-duplex system, a given node \( k \) can transmit, receive or be silent at any given time. Each of these states can be formulated as binary variables as,

\[
r_k = \begin{cases} 
1 & \text{if node } k \text{ receives in the given time slot} \\
0 & \text{otherwise,}
\end{cases}
\]

\[
t_k = \begin{cases} 
1 & \text{if node } k \text{ transmits in the given time slot} \\
0 & \text{otherwise,}
\end{cases}
\]

\[
s_k = \begin{cases} 
1 & \text{if node } k \text{ is silent in the given time slot} \\
0 & \text{otherwise,}
\end{cases}
\]

where a node can be only in one state in a given time slot, which can be written as, \( r_k + t_k + s_k = 1 \), \( \forall k \) or equivalently as \( r_k + t_k \in \{0,1\} \), \( \forall k \). Hence, if \( r_k + t_k = 0 \), it means that node \( k \) is silent in the given time slot.

Next, we define state vectors that consist of binary state variables for each node in the network as,

\[
r = [r_1, r_2, ..., r_K],
\]

\[
t = [t_1, t_2, ..., t_K],
\]

\[
s = [s_1, s_2, ..., s_K].
\]

Based on the above notation, SINR at node \( k \) in a flexible half-duplex network can be expressed as [11],

\[
\text{SINR}_k = r_k \frac{P t d_k}{1 + P t i_k},
\]

where \( P \) denotes the transmit power of each node, and \( d_k \) and \( i_k \) represents the normalized magnitude-squared gains of the desired and interfering fading channels of node \( k \), respectively.

Note that a node operating in flexible-half duplex mode has the choice of transmitting, receiving or being silent. Thus, the variables \( t \) and \( r_k \) in (1) can be manipulated to maximize the sum-rate of the network. Therefore, we devise an optimization problem to maximize the sum-rate of the flexible half-duplex network using (1) as,

\[
\max_{t,r} \sum_{k=1}^{K} \log \left( 1 + r_k \frac{P t d_k}{1 + P t i_k} \right)
\]

Subject to

\[
C1 : t_k, r_k \in \{0,1\}, \forall k,
\]

\[
C2 : t_k + r_k \in \{0,1\}, \forall k.
\]

Even though we consider instantaneous sum-rate in the above optimization problem, this is similar to the average sum-rate optimization in a network where the sum-rate optimization is independent from one time slot to the other.

Please note that, in the absence of any other constraints, treating a silent node as a receiving node would not decrease the sum-rate of the network. This is because, if a node is to receive from a silent node, its received SINR would be zero. This is the same as the SINR of a node that is trying receive from another receiving node. Also, the transmit power of a silent node is zero which is similar to the transmit power of a receiving node. Thus, by removing the second constraint, we set each node to either transmit or receive. After obtaining the optimum state vectors, we can obtain the silent state vector \( s \) using the receiving state vector \( r \) and the desired channel gain matrix \( D \) by redefining any receiving node with zero
The SINR as a silent node. As such, the simplified optimization problem can be given as,
\[
\max_t \sum_{k=1}^{K} \log \left(1 + \left(1 - t_k\right) \frac{P t d_k}{1 + P t i_k}\right)
\]
Subject to
\[
C1 : t_k \in \{0, 1\}, \forall k.
\]

We note that solving (3) is hard due to its non-convexity/non-linearity and the combinatorial nature. Even when we relax the binary constraints to produce
\[
\max_t \sum_{k=1}^{K} \log \left(1 + \left(1 - t_k\right) \frac{P t d_k}{1 + P t i_k}\right)
\]
Subject to
\[
C1 : 0 \leq t_k \leq 1, \forall k,
\]
the relaxed problem in (4) is still non-convex and non-linear. It is possible to obtain the optimal state of the nodes in each time slot through an exhaustive search. However, for a network comprised of \(K\) nodes, there are \(2^K - 1\) possible combinations to search through and the computational complexity is quite high for large \(K\). Therefore, we are motivated to seek an alternative approach to find the optimal states of the nodes.

IV. USE OF APPROXIMATION TECHNIQUES

In this section, we consider the use of two approximation techniques to either linearise or to convexify the simplified optimization problem given in (3) or the relaxed problem given in (4).

A. Tight Lower Bound Approximation

Under the tight lower bound approximation, we relax the non-convex problem using the following lower bound [12]
\[
a \log(z) + b \leq \log(1 + z),
\]
that is tight at \(z = z_0\), when the approximation constants are chosen as
\[
a = \frac{z_0}{1 + z_0}, \quad b = \log(1 + z_0) - \frac{z_0}{1 + z_0} \log(z_0).
\]
Based on this approximation, we can rewrite the optimization problem in (4) as
\[
\max_t \sum_{k=1}^{K} a_k \log \left(1 + \left(1 - t_k\right) \frac{P t d_k}{1 + P t i_k}\right) + b_k
\]
Subject to
\[
C1 : 0 \leq t_k \leq 1, \forall k,
\]
where
\[
a_k = \frac{z_0}{1 + z_0}, \quad b_k = \log(1 + z_0) - \frac{z_0}{1 + z_0} \log(z_0),
\]
\[
z_0 = \left(1 - \bar{t}_k\right) \frac{P \bar{e} d_k}{1 + P \bar{e} i_k}.
\]
and \(\bar{e}\) is the initial solution or the solution achieved via previous iteration.

To convert this non-convex objective function into a convex function we use variable transformation, \(t_k = e^{y_k}\) and \(0 \leq t_k \leq 1 \Rightarrow y_k \leq 0\). The reformulated optimization problem is given by
\[
\max_y \sum_{k=1}^{K} a_k \log(P d_{i,k}) + b_k - a_k \log \left(1 + P \sum_{j=1}^{K} i_{j,k} e^{y_j}\right)
\]
\[
\quad + a_k y_k + a_k \log(1 - e^{y_k})
\]
Subject to
\[
C1 : y \leq 0,
\]
where node \(k\) considers the transmitted signal from node \(l_k\) as desired.

Note that the reformulated problem (5) is concave and could be solved using any convex optimizer for a global solution. Then the binary solution can be obtained using the branch and bound method. However, the complexity of the branch and bound method increases exponentially with the number of nodes in the network. If the maximum number of iterations is limited, the accuracy of the approximation drops when branch and bound method cannot converge within the maximum iteration limit.

B. Arithmetic-Geometric Inequality

According to the arithmetic-geometric inequality given by [13], the following holds.
\[
\sum_{k=1}^{K} x_k^2 \geq \left(\prod_{k=1}^{K} x_k\right)^{1/K},
\]
where \(x_1, x_2, \ldots, x_K\) are non-negative real numbers.

Based on the above inequality, along with the monotonically increasing nature of the log function, we obtain an approximated optimization problem given by
\[
\max_t \sum_{k=1}^{K} (1 - t_k) \frac{P t d_k}{1 + P t i_k}
\]
Subject to
\[
C1 : t_k \in \{0, 1\}, \forall k.
\]
Next, we use 0-1 programming [14], to reformulate this non-linear, non-convex binary fractional optimization problem into a non-convex linear mixed integer optimization problem given by
\[
\max_{k,g,x} \sum_{k=1}^{K} g_k
\]
Subject to
\[
C1 : g_k + P \sum_{j=1}^{K} v_{k,j} i_{j,k} = P \sum_{j=1}^{K} t_j d_{j,k} - P \sum_{j=1}^{K} z_{k,j} d_{j,k}, \forall k,
\]
\[
C2 : v_{k,j} \geq g_k t_j, \forall k, j,
\]
\[
C3 : v_{k,j} \geq g_k + g_k t_j - g_k, \forall k, j,
\]
\[
C4 : z_{k,j} \geq 0, \forall k, j,
\]
\[
C5 : z_{k,j} \geq t_k + t_j - 1, \forall k, j,
\]
\[
C6 : t_k \in \{0, 1\}, \forall k,
\]
where, \( g_k^j = 0 \) and \( g_k^u = \sum_{j=1}^{K} P d_{j,k} \).

Solving the reformulated problem in (6) can be done using mixed integer optimization tools such as LINDO and Matlab “initlinprog” solver or any other linear solver by relaxing the binary constraints.

We note that the approximation provided in (6) is similar to the optimization problem considered in [11]. However, the solution method is different to the algorithm proposed in [11], which consists of an inner loop and an outer loop.

The complexity of the arithmetic-geometric approximation is only \( 2K + K^2 \). However, the accuracy of this approximation depends on the tightness of the inequality. As this upper bound would be tightened only under the equality condition, when \( 1 + \frac{\rho t d_{j,k}}{1 + \rho t d_{j,k}} \) values have high deviation, this approximation becomes less accurate.

V. PROPOSED PATTERN SEARCH ALGORITHM

In this section, we focus on developing a low-cost algorithm to solve the optimization problem in (3) without the use of approximations or relaxations. As derivative based algorithms cannot be used directly in discrete optimization, we consider the pattern search method that uses direct search based on the function value [15]. More specifically, we use the exploratory pattern search that looks for an improving direction in which to move from among a set of nearby points. Even though this method was originally developed for continuous optimization problems, it can be used for our problem by selecting \( K \) neighbouring points through changing each of the \( K \) elements in vector \( t \). In each step we will analyse the element that needs to be changed to have the highest increase in the function value and vector \( t \) would be updated accordingly. We will continue the search of improving direction until it converges to a solution.

It is important to note that our optimization problem is non-convex. As such, the pattern search method would only converge to a local solution and this converged solution depends on the starting point of the algorithm. Thus, we further modify our proposed pattern search algorithm to consider multiple starting points. Through simulations we realize that it is sufficient to consider only \( K_{Tx} \) starting points, where \( K_{Tx} \) is the number of total transmitting nodes, so that the resulted set of local solutions would contain the global solution. We select those starting points that at each starting point, only one transmitting node would be active as given by Algorithm 1.

Even though we cannot guarantee global convergence, extensive simulations indicate that it nearly always finds the global solution achieved through exhaustive search. In addition, numerical examples show that the proposed algorithm performs better than alternative resource allocation methods. We note that the complexity of our proposed iterative algorithm is \( K^3 + K^2 \), which is significantly better compared to the exponential nature of exhaustive search.

VI. NUMERICAL RESULTS

In this section, we present numerical examples to illustrate the performance of our proposed techniques.

Algorithm 1: Proposed Pattern Search Algorithm

<table>
<thead>
<tr>
<th>Input</th>
<th>Transmitting nodes list ( K_{Tx} ) and desired and interference channel fading matrices: ( D ), ( I )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>Maximum sum-rate: ( R_{opt} ) and optimum node status vector: ( t_{opt} )</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
&\text{1. } t_{opt} \leftarrow 0 \\
&\text{2. } \text{for } j=\text{each node in } K_{Tx} \text{ do} \\
&\quad t_{pattern} \leftarrow \text{starting point with only node } j \text{ active} \\
&\quad k \leftarrow j \\
&\quad \text{while true do} \\
&\quad \quad \text{for } i=\text{each node in } K_{Tx} \text{ do} \\
&\quad \quad \quad t = t_{pattern}, \quad t(i) \leftarrow 1 \\
&\quad \quad \quad R \leftarrow \text{sum-rate for } t \\
&\quad \quad \quad \text{end} \\
&\quad \quad \text{index} \leftarrow \text{index of maximum rate from } R \\
&\quad \quad \text{if } \text{index} == k \text{ then} \\
&\quad \quad \quad \text{break} \\
&\quad \quad \quad \text{else} \\
&\quad \quad \quad \quad t_{pattern}(\text{index}) \leftarrow 1 \\
&\quad \quad \quad \quad k \leftarrow \text{index} \\
&\quad \quad \text{end} \\
&\quad \text{R}_{max}(j) \leftarrow \text{sum-rate for } t_{pattern} \\
&\quad \text{t}_{max}(j) \leftarrow t_{pattern} \\
&\quad \text{end} \\
&\text{3. } R_{opt} \leftarrow \text{maximum rate from } R_{max} \\
&\text{4. } t_{opt} \leftarrow \text{relevant } t_{max} \\
\end{align*}
\]

A. General network with user pairs

We consider a network with \( D_A \times D_A \) m² coverage and divide it into smaller cells. We place one base station (BS) at the centre of each cell and one user equipment (UE) at a random location within the cell. We assume that the links between the user and the base station in the same cell are desired and that every other link causes interference.

In addition, we assume that the channels between transmitting and receiving nodes have a Rayleigh distribution and the mean of the channel gain is calculated using the standard path-loss model [16] as,

\[
E[h_L]^2 = \left( \frac{c}{4\pi f_c} \right)^2 d_L^{-\beta},
\]

where \( c \) is the speed of light, \( f_c \) is the carrier frequency, \( d_L \) is the distance between the transmitter and the receiver of link \( L \), \( \beta \) is the path loss exponent and \( h_L \) is the instantaneous channel fading gain of link \( L \). For this example, \( f_c = 1.9 \) GHz, \( \beta = 3.5 \) and transmit bandwidth is 200 kHz.

First, we consider the performance accuracy of the two approximation techniques discussed in Section IV and compare with the proposed pattern search algorithm in Section V.

Fig. 1 shows the change of sum-rate versus average received signal-to-noise-ratio (SNR). We observe that, in the low SNR regime the arithmetic mean-geometric mean approximation in

\[1\] Due to the large size of the network, it was not possible to implement exhaustive search.
To allow each cell to select uplink or downlink to optimize its own rate irrespective of other cells. These results can be further explained by the type of network we considered, which would provide similar rate throughput for uplink and downlink when there is no interference. If we consider the frequency reuse scheme, we can see closer performance to conventional TDD scheme at denser area while frequency reuse become underutilization of resources when cells are too far apart. Therefore, we can conclude that the proposed scheme of optimizing over sum-rate actually improves the overall network throughput in denser networks where cells are too close to each other.

Then, we look into the fairness of the proposed scheme with respect to uplink/downlink rate throughputs in each cell. Fig. 3 shows that the percentage of throughput distribution among uplink and downlink for the proposed scheme is similar to that of conventional TDD scheme. We understand that these results are limited to networks where desired user pairs have uniformly distributed distances as considered in this paper and that the fairness can suffer in a non-uniformly distributed network where some desired pairs are far apart from each other compared to other pairs. Therefore, we propose sum-rate optimization with proportional fair scheduling to be considered in future work to overcome this unfairness arises from non-uniform distribution of desired distances.
Fig. 4: sum-rate vs. transmit power for a simple D2D network

B. Simple cellular network with a D2D pair

Next, we consider an example of a traditional half-duplex cellular network to show how remodelling as a flexible network would increase sum-rate performance. Most specifically, we consider a network of one base station and two cellular users where one user is operating in the uplink and other user is operating in the downlink. The network also include two other users that operates as a half-duplex D2D pair. We assume Rayleigh fading with mean calculated using the standard path-loss model given by (7).

We consider the exhaustive search and two benchmark schemes currently used in D2D communication: mode selection between D2D communication and cellular communication [17] and interference aware resource allocation [18]. Under the first benchmarking scheme, we only let either cellular or D2D communication to take place based on the method that has the highest average throughput. In the second scheme, we let the best link in cellular communication to be active as well as the best link in D2D pair to be active causing interference to each other.

Fig. 4 illustrates the change of sum-rate of the network with respect to transmit power. We can see that the proposed technique has better performance than both reference techniques due to its ability to combine both those aspects through flexible duplex remodelling. It also shows that the proposed pattern search algorithm has similar results to exhaustive search.

VII. CONCLUSION

Centralized link scheduling in flexible half-duplex networks was analysed in terms of instantaneous sum-rate optimization. A new optimization problem was derived for a flexible half-duplex network where nodes can choose to transmit, receive or be silent at a given time. The new optimization problem was solved using two approximation techniques as well as a novel low-cost pattern search algorithm. Extensive numerical examples were used to illustrate the performance of our proposed algorithm. In addition, it was shown that the flexible half-duplex technique can combine mode selection and interference aware resource allocation in D2D communication underlying cellular network to improve overall network performance. As we consider the instantaneous sum-rate, any change in the network, including the number of nodes, SINR or node locations would require re-computation of the optimum state for each node.

REFERENCES

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Dayarathna, S; Razlighi, M; Senanayake, R; Zlatanov, N; Evans, J

Title:
Centralized Scheduling with Sum-Rate optimization in Flexible Half-Duplex Networks

Date:
2020-05

Citation:

Persistent Link:
http://hdl.handle.net/11343/241462