On the minimum latency transmission scheduling in wireless networks with power control under SINR constraints†


Automatic Control Lab, School of Electrical Engineering, Royal Institute of Technology (KTH), Osquldas väg 10, 100-44 Stockholm, Sweden

ABSTRACT

In order to alleviate interference and contention in a wireless network, we may exploit the existence of multiple orthogonal channels or time slots, thus achieving a substantial improvement in performance. In this paper, we study a joint transmission scheduling and power control problem that arises in wireless networks. The goal is to assign channels (or time slots) and transmitting powers to communication links such that all communication requests are processed correctly, specified quality-of-service requirements are met, and the number of required time slots is minimised. First, we formulate the problem as a mixed-integer linear programming. Then, we show that the problem considered is non-deterministic polynomial-time hard, and subsequently, we propose non-trivial bounding techniques to solve it. Optimisation methods are also discussed, including a column generation approach, specifically designed to find bounds for the transmission scheduling problem. Moreover, we develop optimisation techniques in which the bounding techniques are integrated in order to derive the optimal solution to the problem faster. We close with an extensive computational study, which shows that despite the complexity of the problem, the proposed methodology scales to problems of non-trivial size. Our algorithms can therefore be used for static wireless networks where propagation conditions are almost constant and a centralised agent is available (e.g. cellular networks where the base station can act as a centralised agent or wireless mesh networks), and they can also serve as a benchmark for the performance evaluation of heuristic, approximation or distributed algorithms that aim to find near-optimal solutions without information about the whole network. Copyright © 2013 John Wiley & Sons, Ltd.

*Correspondence
T. Charalambous, Automatic Control Lab, School of Electrical Engineering, Royal Institute of Technology (KTH), Osquldas väg 10, 100-44 Stockholm, Sweden.
E-mail: themisc@kth.se
†Preliminary results of this work have been published as a Technical Report in [1].

Received 17 October 2012; Revised 27 November 2012; Accepted 14 December 2012

1. INTRODUCTION

Wireless technology standards provide a radio-frequency spectrum with a set of many non-overlapping channels, and a node has the option to choose on which channel to transmit. Likewise, in cases where only a single channel is available, it is possible to divide time into frames, and then, frames can be divided into time slots, such that at each frame a node has the option to choose on which time slot to transmit. In the latter case, synchronisation of the wireless nodes in the network is necessary. If synchronisation is not considered, however, choosing a channel or a time slot in the network becomes the same problem. It is important to schedule channel/time slot access in such a way that spatial reuse is fully exploited, and hence, the number of channels required to successfully complete all requests is minimised.

Power control has been a prominent research area with increased interest (e.g. [2–9]). Increased power ensures longer transmission distance and higher data transfer rate. However, power minimisation increases not only battery lifetime but also the effective interference mitigation that increases the overall network capacity by allowing higher frequency reuse. Power control has been extensively employed for medium access control in multi-hop wireless networks (e.g. [10–15]). Some of them aimed
to minimise power dissipation. For example, ElBatt and Ephremides [16] proposed a two-phase method for the joint scheduling and power control, which aims to find an admissible set of links along with their transmission power levels in a single channel only. Behzad and Rubin [17] studied the same problem as in [16], but it focuses on minimising the scheduling length. Others, such as Muqattash and Krunz [18] and Tang et al. [11], aimed to maximise throughput at the cost of increased power dissipation by allowing many simultaneous interference-limited transmissions. In such schemes, either time is divided into fixed-length slots or there exist many channels, and the wireless nodes have to choose on which one to transmit. The minimum latency scheduling problem (e.g. [19–21] and references therein) and the computation of efficient schedules for the abstract physical model with power control (e.g. [10, 11, 14–16, 22–24]) have been both extensively studied. However, many approaches concentrate, for example, on throughput maximisation in a single or multiple channels (e.g. [25, 26]) and scheduling length minimisation (e.g. [27, 28] and reference therein), rather than the minimisation of the number of channels or time slots required. A wide range of applications for wireless networks are time-critical and imposes stringent requirements on the communication latency. For example, a given rate demand is requested (that should be satisfied with minimum-length periodic scheduling actions), or a given volume of traffic must be delivered to the destinations in minimum time. However, minimising the scheduling length requires coordination between the wireless nodes in order to orchestrate the order, duration and initialisation of transmissions, something which introduces extra communication overheads. On the other hand, minimum latency transmission scheduling (MLTS) does not require the communication between the nodes, but simply synchronisation to a global clock in the network, so that the nodes are able to know the beginning and the end of slots.

When studying wireless networks, the choice of model is crucial. Not only must the chosen model facilitate the design of protocols but it also has to truthfully reflect the nature of the real network. Fading-channel models depict real-world phenomena in wireless communications. These phenomena include multi-path fading, shadowing and attenuation with distance. Although fading effects have been considered as detrimental in 2G wireless networks, in 3G networks, they are seen as an opportunity to increase the capacity that incorporate data traffic [29]. The most common fading-channel model being used is the physical model, which is thoroughly described in Section 2. On a finer granularity, one distinguishes between the geometric and the abstract physical model. In the geometric physical model, the channel gain between two nodes is solely determined by their spatial distance. Hence, simplifying assumptions are incorporated into this model; for example, the radios are perfectly isotropic, and there are no obstructions [30]. In the abstract physical model, the channel gain between two nodes incorporates all the real-world phenomena, and hence, no information can be extracted about the geometry of the network.

The rest of the section concentrates on the related work that considers transmission scheduling under the (geometric and abstract) physical model only. For the geometric physical model, the non-deterministic polynomial-time (NP) hardness of wireless scheduling without power control is proven in [30]. In [31], it is proven that strategies in the geometric physical model that use uniform power-assignment schemes (same power to all nodes in the network) or linear power-assignment schemes (power levels proportional to the minimum power required to reach the receiver node) have a bad scheduling complexity. In addition, they propose a power-assignment algorithm that successfully schedules a network using a poly-logarithmic number of time slots. Approximation scheduling algorithms (e.g. [32, 33]) are proposed that compute a feasible solution in polynomial time for the geometric model with worst-case approximation that guarantees for arbitrary network topologies when the power levels are constant. In [34], it is shown that solutions with oblivious power assignments (the power level of a node depends only on the transmitter-receiver distance) cannot compete with solutions using possibly different power levels and channels for a network. However, they are capable of achieving nearly the same performance as solutions restricted to symmetric power and channel assignments. For the geometric physical model, the NP-hardness of wireless scheduling without power control is proven in [30] and with power control is proven in [35], given that we know the minimum \( P_{\text{min}} > 0 \) and maximum \( P_{\text{max}} < \infty \) transmission power levels.

When considering the abstract physical model, less information has to be considered, and the values in the gain matrix are not restricted by the topology of the network (not every gain matrix of the abstract model can be expressed as a network, and on the contrary, every gain matrix of the geometric model could be a case for the abstract model); that is, in the geometric physical model, we have the advantage of exploiting the geometry of the network in order to check complexity and to design scheduling algorithms. In [36], the NP-hardness of wireless scheduling without power for the abstract physical model is proven and the problem is analytically solved via a column generation (CG) approach. However, in order to achieve the best performance, scheduling and power control should be optimised jointly. This problem is notoriously difficult to solve, even in a centralised manner. In [37], this line of research is followed and the transmission scheduling problem for minimising the total number of slots (channel or time) for variable power levels is formulated. However, in this work, they considered non-problem-specific analytic solutions that are computationally expensive.

We consider the abstract physical model, and the contributions of this paper can be summarised as follows.

(1) We first prove that the MLTS problem for the abstract model is NP-hard for variable power
levels. Contrary to existing approaches and results in the literature, our formulation includes the choice of optimal transmitting powers and arbitrary topology. Furthermore, the generality of the abstract physical model implies the complexity of the geometric physical model and completes the theory that the transmission scheduling problem with power control for the physical model in general is NP-hard. This answers the open problem posted in [38].

2. MODEL

The system model can be divided into two levels: the network as a whole and the channel. Thus, we have the network model and the channel model. The network model concerns the general topology of the nodes and their characteristics. The channel model describes the assessment of the link quality between communication pairs and the interaction between the nodes in the network.

2.1. Network model

In this study, we consider a network where the links are assumed to be unidirectional and each node is supported by an omnidirectional antenna. For a planar network (easier to visualise without loss of generality), this can be represented by a graph \( \mathcal{G} = (\mathcal{N}, \mathcal{L}) \), where \( \mathcal{N} \) is the set of all nodes and \( \mathcal{L} \) is the set of the active links in the network. Each node can be a receiver or a transmitter only at each time instant because of the half-duplex nature of the wireless transceiver. Each transmitter aims to communicate with a single node (receiver) only, which cannot receive from more than one node simultaneously. We denote by \( \mathcal{T} \) the set of transmitters and \( \mathcal{R} \) the set of receivers in the network.

2.2. Channel model

The link quality is measured by the signal-to-interference-and-noise ratio (SINR). The channel gain on the link between transmitter \( i \) and receiver \( j \) is denoted by \( g_{ij} \) and incorporates the mean path-loss as a function of distance, shadowing and fading, as well as cross-correlations between signature sequences. All the \( g_{ij} \)'s are positive and can take values in the range \( [0, 1] \). Without loss of generality, we assume that the intended receiver of transmitter \( i \) is also indexed by \( i \). The power level chosen by transmitter \( i \) is denoted by \( p_i \). \( v_i \) denotes the variance of thermal noise at the receiver \( i \), which is assumed to be additive Gaussian noise. The interference power at the \( i \) node, \( I_i \), includes the interference from all the transmitters in the network and the thermal noise, and is given by

\[
I_i = \sum_{j \neq i, j \in \mathcal{T}} g_{ji} p_j + v_i \quad (1)
\]

Therefore, the SINR at the receiver \( i \) is given by

\[
\Gamma_i = \frac{g_{ii} p_i}{\sum_{j \neq i, j \in \mathcal{T}} g_{ji} p_j + v_i} \quad (2)
\]

The quality of service (QoS) is measured in terms of SINR. Hence, independent of nodal distribution and traffic pattern, a transmission from transmitter \( i \) to its corresponding receiver is successful (error-free) if the SINR of the receiver is greater or equal to the capture ratio \( \gamma_i \). The value of \( \gamma_i \) depends on the modulation and coding characteristics of the radio. Therefore, we require that

\[
\frac{g_{ii} p_i}{\sum_{j \neq i, j \in \mathcal{T}} g_{ji} p_j + v_i} \geq \gamma_i \quad (3)
\]

3. PROBLEM FORMULATION

In this section, we present the problem of finding the minimum possible number of time slots (or channels) and
the corresponding transmitting powers, such that all communication requests are being processed correctly and QoS requirements for successful transmissions are satisfied.

Note that to ensure feasibility of our problem, we can define the deadline of the network as follows. The maximum number of time slots that may be required is equal to the number of links, $|\mathcal{L}|$. Henceforth, we will assume that the first time slot is at time 1; the latest point in time for which there can be a scheduled transmission is therefore $D = |\mathcal{L}|$. The notation used for the networks in this paper is given in Notation 1.

**Notation 1** Notation used for the networks:

| $\mathcal{N}$ | The set of all nodes in the network |
| $\mathcal{T}$ | The set of transmitters in the network |
| $\mathcal{R}$ | The set of receivers in the network |
| $g_{ij}$ | The channel gain on the link $i \rightarrow j$ |
| $v_i$ | The variance of thermal noise at the receiver $i$ |
| $I_i$ | The interference power at the $i$th receiver |
| $\Gamma_i$ | The SINR at the $i$th receiver |
| $\gamma_i$ | The capture ratio at the $i$th receiver |
| $D$ | The deadline of the network |

To formulate the optimisation problem, we define two sets of decision variables, for each transmitter $i \in \mathcal{T}$ and time $t = 1, \ldots, D$; processing-time variables:

$$x_i(t) = \begin{cases} 
1, & \text{if transmitter } i \text{ is active at time } t \\
0, & \text{otherwise} 
\end{cases}$$

(4)

and power level variables: $p_i(t) \in \mathbb{R}_+^+$. Because the problem involves both integer and continuous decision variables, the mathematical formulation is classified as a mixed-integer program and is given in Model 1.

Objective (5a) minimises the number of time slots needed to schedule all the transmitters in the network. For every transmitter $i \in \mathcal{T}$, $\sum_{t=1}^{D} x_i(t)$ is equal to the scheduling time, because only one of $x_i(1)$, $x_i(2), \ldots, x_i(D)$ will be equal to 1. We define as $\tau$ the latest scheduling time for a transmission. We can linearise the objective by requiring $\tau$ to be larger than or equal to all of the pairs’ scheduling time, that is, $\tau \geq \sum_{t=1}^{D} x_i(t)$ for all $i \in \mathcal{T}$. Constraint (5b) ensures that each link in the network is processed at least once in the schedule. Note that there is always an optimal schedule in which each link is processed only once. Constraint (5c) makes sure that if a pair is not processed at a specific time slot, then the power level of the corresponding transmitter is 0 at that time slot. The QoS conditions are guaranteed by constraint (5d). The constraint only affects the optimisation if $x_i(t)$ takes the value 1. Finally, the last two constraints (5e) and (5f) define the admissible values for the decision variables.

### Model 1

| Model 1 Minimum number of time slots |

| minimise $\tau = \max_{i \in \mathcal{T}} \sum_{t=1}^{D} t x_i(t)$ |
| subject to |
| $\sum_{t=0}^{D} x_i(t) \geq 1 \ \forall i \in \mathcal{T}$ |
| $x_i(t) = 0 \Rightarrow p_i(t) = 0 \ \forall i \in \mathcal{T}, \ t = 1, \ldots, D$ |
| $x_i(t) = 1 \Rightarrow g_{ii} p_i(t) \geq \gamma_i \left( \sum_{j \in \mathcal{T}, j \neq i} g_{ji} p_j(t) + v_i \right)$ |
| $\forall i \in \mathcal{T}, \ t = 1, \ldots, D$ |
| $x_i(t) \in \{0, 1\} \ \forall i \in \mathcal{T}, \ t = 1, \ldots, D$ |
| $p_i(t) \in \mathbb{R}_+^+, \forall i \in \mathcal{T}, \ t = 1, \ldots, D$ |

### 4. PRELIMINARIES

Inequality (3) depicts the QoS requirement of a communication pair $i$ while transmission takes place. After manipulation it becomes equivalent to the following:

$$p_i \geq \gamma_i \left( \sum_{j \neq i} g_{ji} p_j + \frac{v_i}{\gamma_{ii}} \right)$$

(6)

In matrix form, for a network consisting of $n$ communication pairs, this can be written as

$$p \geq \Gamma G p + \eta$$

(7)

where $\Gamma = \text{diag} (\gamma_i), p = (p_1 \ p_2 \ \ldots \ p_n)^T, \eta_i = \frac{v_i}{\gamma_{ii}}$ and

$$G_{ij} = \begin{cases} 
0, & \text{if } i = j \\
\frac{g_{ji}}{g_{ii}}, & \text{if } i \neq j 
\end{cases}$$

Let

$$C = \Gamma G$$

(8)

so that (7) can be written as

$$(I - C)p \geq \eta$$

(9)

Matrix $C$ has strictly positive off-diagonal elements, which implies that it is irreducible, because we are not considering totally isolated groups of links that do not interact with each other. By the Perron–Frobenius theorem [39], we have that the spectral radius of the matrix $C$ is a simple eigenvalue, whereas the corresponding eigenvector is positive component-wise. A necessary and sufficient condition
for the existence of a nonnegative solution to inequality (9) for every positive vector \( \eta \) is that \((I - C)^{-1}\) exists and is nonnegative. However, \((I - C)^{-1} \geq 0\) if and only if \( \rho(C) < 1\) [39] (where \( \rho(C) \) denotes the spectral radius of \( C \)), or, equivalently, \((C - I)\) is Hurwitz (because \((C - I)\) is Metzler) [40].

Therefore, the necessary and sufficient condition for Equation (7) to have a positive solution \( p^* \) for a positive vector \( \eta \) (i.e. there exists a set of powers such that all the senders can transmit simultaneously and still meet their QoS requirements (the minimum SINR for successful reception)) is that the Perron–Frobenius eigenvalue of the matrix \( C \) is less than 1.

5. COMPUTATIONAL COMPLEXITY

Theorem 1. Problem (5) is NP-hard.

Proof. The statement of Theorem 1 is equivalent to saying: deciding whether the optimal value of Equation (5) does not exceed a given value \( \kappa \in \mathbb{N} \) is NP-hard. We construct a polynomial-time reduction of the graph colouring problem, which is well known to be NP-hard [41]. Given an undirected graph \( G = (V, E) \) with nodes \( V = \{1, \ldots, n\} \) and edges

\[
E \subseteq \{\{i, j\} : i, j \in V, i \neq j\}
\]

as well as a scalar \( k \in \mathbb{N} \), the graph colouring problem asks whether there is an assignment \( f : V \mapsto \{1, \ldots, k\} \) of nodes to \( k \) colours such that \( f(i) \neq f(j) \) for all \( \{i, j\} \in E \); that is, neighbouring nodes must have different colours. Our reduction takes as input a graph colouring instance and generates an instance of Equation (5) such that the optimal value of problem (5) does not exceed \( \kappa = k \) if and only if the answer to the graph colouring problem is affirmative. Towards this end, we set \( \mathcal{F} := V, D := |V| = n, \mathcal{L} := E, y = (1, \ldots, 1)^T, v_{ij} = 0 \forall i \in \mathcal{F} \) and

\[
g_{ij} := \begin{cases} 1 & \text{if } \{i, j\} \in E \\ 1/2 & \text{if } i = j \\ 1/(2n) & \text{otherwise} \end{cases}
\]

The size of this reduction is polynomial in the size of the graph colouring instance. Hence, if we show that the optimal value of Equation (5) does not exceed \( \kappa \) if and only if the answer to the graph colouring instance is affirmative, we have proven that the solution of Equation (5) is NP-hard. We proceed in two steps. Firstly, we show that if there is a graph colouring that uses \( \zeta \) colours, then the optimal value of Equation (5) is smaller or equal to \( \zeta \). Secondly, we show that if there is a feasible solution for Equation (5) of value \( \zeta \), then we can construct an admissible \( \zeta \)-colouring for the graph colouring instance.

The assertion follows from the combination of both arguments and the fact that we consider a minimisation objective.

As for the first step, assume that there exists a colouring \( f : V \mapsto \{1, \ldots, \zeta\} \). Given this colouring, we construct a feasible solution \((x, p)\) of objective value \( \zeta \). Towards this end, we set \( x_i(t) := 1 \) if \( \hat{f}(i) = t \) and \( x_i(t) := 0 \) otherwise, \( \forall i \in \mathcal{F} \). Likewise, set \( p_t := p_t(\hat{p}_t \in \mathbb{R}_+) \) if \( \hat{f}(i) = t \) and \( p_t := 0 \) otherwise, for all \( i \in \mathcal{F} \) and \( t = 1, \ldots, D \). By construction, constraints (5b), (5c), (5e) and (5f) are satisfied, for any value \( \hat{p}_t, \hat{p}_t \in \mathbb{R}_+ \). For \( i \in \mathcal{F} \) and \( t \in \{1, \ldots, D\} \) with \( x_i(t) = 1 \), constraint (5d) requires that

\[
p_t(t) \geq 2 \sum_{\{i,j\} \in E} p_{j}(t) + \frac{1}{n} \sum_{j \in V, j \neq i} \sum_{\{i,j\} \notin E} p_j(t)
\]

Because \( f \) constitutes a valid colouring, the first term on the right-hand side must evaluate to zero, because otherwise, the spectral radius \( \rho \) of the matrix that is constituted by \( \{i, j\} \in E \) is greater than 1 and hence the network would be infeasible, as described in Section 4. On the other hand, when the first term of the right-hand side is zero, the second term fulfills the inequality because

\[
\sum_{j \in V, j \neq i} \frac{1}{n} < 1 \Rightarrow \|C\|_\infty < 1
\]

and hence, \( \rho(C) < 1 \), where \( C \) consists of \( i, j \in V, j \neq i, \{i, j\} \notin E \). We conclude that constraint (5d) is also satisfied by our choice of \( x \) and \( p \in \mathbb{R}_+ \). Note that the objective function (5a) evaluates to \( \zeta \) for the constructed solution \((x, p)\), which implies that the optimal value of Equation (5) must be smaller or equal to \( \zeta \).

In the second step, we use a feasible solution to Equation (5) with objective value \( \zeta \) to construct a valid colouring of the graph \( G \) with at most \( \zeta \) colours. Assume that we have a feasible solution \((x, p)\) for problem (5) with objective value \( \zeta \). Because we consider a minimisation problem, without loss of generality, we can assume that \( \sum_{t=1}^{D} x_i(t) = 1 \). Hence, we obtain a function if we set \( \hat{f}(i) := t \) if and only if \( x_i(t) = 1 \). Furthermore, because the objective value (5a) of \((x, p)\) is \( \zeta \), the range of \( f \) is limited to \( \{1, \ldots, \zeta\} \). We now show that \( f \) constitutes a valid colouring of graph \( G \), that is, \( f(i) \neq f(j) \) for all \( \{i, j\} \in E \). Assume to the contrary that there is \( \{i, j\} \in E \) with \( f(i) = f(j) = \hat{t} \). In this case, \((x, p)\) must satisfy the constraints

\[
p_t(\hat{t}) \geq 2 \sum_{\{i,j\} \in E} p_{j}(\hat{t}) + \frac{1}{n} \sum_{j \in V, j \neq i} \sum_{\{i,j\} \notin E} p_j(\hat{t}) \geq 2 p_j(\hat{t})
\]
We can check if two nodes can be processed simultaneously by isolating them as a network and by checking its spectral radius. We construct set $CL = \mathcal{N}$ and sort it in decreasing order according to $\sum_{j \in \mathcal{F}} e_{ij}$, that is, the number of the nodes which node $i$ cannot be simultaneously processed with. The first member of set $CL$ is deemed as the first slot leader. Then, each node in position 2 onwards is checked with all the nodes in previous positions. If a node can be simultaneously present in the same time slot with any one of the checked nodes, then the node is removed from set $CL$. Otherwise, it becomes a slot leader. The number of slot leaders at the end of this operation is the obtained lower bound $LB$. Note that because we check if two nodes can be in the same slot in each case and not the whole set of nodes assigned to the specific slot, this scheme will allow more nodes in the slot than it would otherwise admit (the final selection is therefore not necessarily feasible), and hence, less slots will be required in total. That is why this methodology constitutes an LB technique.

We also present a variation of the lower bound $LB$, referred to as $LB'$, which includes two additional steps to Algorithm 1. The first addition to the algorithm is performed at the end of each iteration of the WHILE loop. In particular, we update values $e_{ij}$ for $i, j \in CL$ and consequently update and resort set $CL$. The final addition is a checking process that may increase the lower bound by 1. To explain this, Algorithm 1 implicitly assumes that a number of nodes can be simultaneously executed in the same time slot if they at least can be simultaneously executed with the so-called slot leader. In polynomial time, we can check whether the nodes implicitly assumed to be executed with the final slot leader cannot be pairwise simultaneously executed. If we can find such ‘infeasible’ pairs, then $LB'$ is equal to $LB + 1$. To clarify this, we note that the members of the infeasible pair cannot be both present in the final slot. Furthermore, neither of the members can be scheduled at any previous time slots (by construction of the lower bound). Hence, one of the members of the pair must be scheduled in a new time slot.

### 6.1.2. Upper bound

In the literature, UB techniques are described as approximation scheduling methods (e.g., [32, 33, 35] and references therein). We describe two non-trivial upper bound methods that, when combined with strong LB techniques, are capable of closing the optimality gap, hence finding the optimal solution efficiently. We first describe a simple yet effective heuristic, based on a priority scheduling policy. The derived solution value, referred to as $UB$, serves also as a cut-off value in the B&B approach described in the next subsection. The basic idea of the policy is to keep adding new transmissions at the current time slot according to a priority criterion, until no more transmissions can be scheduled without violating the SINR constraints. In such a case, the next time slot is considered, and the process is repeated for all the remaining unscheduled transmission

```plaintext
Algorithm 1 Lower bounding technique LB

initialise
Set $e_{ij} \leftarrow 1$ if $\rho([0,c_{ij} ; c_{ji} , 0]) > 1$ \forall i, j \in \mathcal{F}$ and set $CL_i = i$. Resort $CL$ in decreasing order of $\sum_{j \in \mathcal{F}} e_{ij}$.

Set $k \leftarrow 2.$

while $k \leq |CL|$ do
    $i \leftarrow CL_k$.
    for $m = 1$ to $k - 1$ do
        $j \leftarrow CL_m$.
        if $\rho([0,c_{ij} ; c_{ji} , 0]) < 1$ then
            Remove $i$ from $CL$.
            $k \leftarrow k - 1$.
            Exit FOR loop.
        end if
    end for
    $k \leftarrow k + 1$
end while
Set $LB \leftarrow |CL|$.
return $LB$
```
pairs. Note that for each node considered for a time slot, the spectral radius of the matrix that constitutes the network is calculated, which takes time $O(n^3)$. In our algorithm, the priority value of pair $i \in \mathcal{T}$ is found using

$$ R_i = \frac{v_i y_i}{g_{ii}} \quad (10) $$

which effectively represents the power that transmitter $i$ produces in a time slot when it is the only active transmitter. The complete heuristic algorithm is described in Algorithm 2. Note that set $A$ (in the description that follows) contains all transmission pairs in decreasing order of their priorities.

**Algorithm 2 Upper bounding technique**

**initialise**

$S = \emptyset$

$A = \{i_1, \ldots, i_k | i_k \in \mathcal{T}, \forall k \in |\mathcal{T}|, R_{ik_1} \geq R_{ik_2}$ if $k_1 \leq k_2\}

UB $ \leftarrow 0$

while $S \neq A$

for $j \in A$

if $j \cup set$ satisfies SINR constraints then

$set \leftarrow set \cup \{j\}$

$A \leftarrow A \setminus \{j\}$

$S \leftarrow S \cup \{j\}$

end if

end for

UB $\leftarrow$ UB $+ 1$

end while

return UB

Note that, in this case, because not the optimal set of pairs is chosen to be admitted in each slot, the allocation of the pairs will be suboptimal and the number of slots required will be an upper bound for the minimum number of slots. For constant noise, the priority value is equivalent to the sorting in [32] of the links by nondecreasing order of length. However, apart from taking into account that thermal noise could differ at the receivers, our approach calculates the spectral radius of the matrix each time a node is admitted in a network, thus allowing for variable power; hence, it is less conservative than the affectedness and affectance used in [32] and [33], respectively, because it is essentially a metric larger than $\|C\|_{\infty}$, which is already conservative [42].

While UB produces very good upper bounds, we propose an additional UB algorithm, called $UB'$ and described in Algorithm 3, that in some instances outperforms UB. This approach chooses the order by which the nodes enter a network by the interference they experience. The algorithm is as follows. The first step of the algorithm is to calculate the sum of each row and column in matrix $C$ as given by Equation (8). The sum of each column represents the interference caused by a particular node, whereas the sum of each row represents the interference experienced by the particular node. The algorithm chooses the node with maximum caused interference and places it in a time slot. The algorithm then checks the remaining nodes in decreasing order of the interference experienced. If a node is feasible with the members in a time slot, then it is included in the same set. Otherwise, it is placed in a new time slot.

**Algorithm 3 Upper bounding technique**

**initialise**

Set $\text{sumrows}_i \leftarrow \sum_{j \in \mathcal{T}} c_{ji}$ and $\text{sumcols}_i \leftarrow \sum_{j \in \mathcal{T}} c_{ij}$.

Find $icol \in \mathcal{T}$ such that $\text{sumcols}_{icol}$ is maximum.

Set $k \leftarrow 1$ and $S_k \leftarrow \{icol\}$. Update $\mathcal{T} \leftarrow \mathcal{T} \setminus \{icol\}$.

while $\mathcal{T} \neq \emptyset$

Find $irow \in \mathcal{T}$ such that $\text{sumrows}_{irow}$ is maximum.

for $m = 1$ to $k$

if $irow \cup S_m$ satisfies SINR constraints then

$S_m \leftarrow S_m \cup \{irow\}$

Exit FOR loop.

end if

end for

if $irow$ could not be placed in any of the sets $S$ then

$k \leftarrow k + 1$, $S_k \leftarrow \{irow\}$

end if

$\mathcal{T} \leftarrow \mathcal{T} \setminus \{irow\}$

end while

Set $UB \leftarrow k$.

return UB

We also present a variant of the UB technique given in Algorithm 3, which includes an additional step. In particular, we add a resorting process at the end of each iteration of the WHILE loop. The process finds the minimum total power emitted by the nodes in each set $S$ such that the SINR constraints are satisfied. The sets are then resorted in increasing order of the total power. The idea is to add more members to the sets with lower power levels first. The resulting upper bound is referred to as $UB''$. Our UB techniques are compared in Section 7 with ApproxLogN algorithm proposed in [32], which is considered the current state-of-the-art.

### 6.1.3. Column generation method.

We also describe a CG technique, based on an alternative formulation of the original problem, which is capable of providing both a lower bound and an upper bound. It is important to note that the main contribution of the method is the fact that it obtains stronger lower bounds. The new formulation uses an explicit representation of feasible sets of transmission pairs. A set of transmission pairs...
s \subseteq \mathcal{I} is said to be feasible if the simultaneous execution of all the pairs in the set does not violate the SINR constraint (5d).

The new set covering formulation, given in Model 2, incorporates the complete set of feasible sets of transmission pairs, S. A binary decision variable is associated to each feasible sequence s ∈ S, defined as

\[ \theta_s = \begin{cases} 
1, & \text{if set } s \text{ is used in the optimal solution} \\
0, & \text{otherwise}
\end{cases} \]

**Model 2** Set covering formulation

\[
\begin{align*}
\text{minimise} & \quad \sum_{s \in S} \theta_s \quad (11a) \\
\text{subject to} & \quad \sum_{s \in \mathcal{I} \subseteq S} \theta_s \geq 1 \quad \forall i \in \mathcal{I} \quad (11b) \\
& \quad \theta_s \in \{0, 1\} \quad \forall s \in S \quad (11c)
\end{align*}
\]

The objective is to minimise the number of sets that are required in the optimal solution. Constraint (11b) ensures that the solution includes at least one set for each pair i ∈ \mathcal{I}, and constraint (11c) defines the allowable range of values for the decision variables.

In order to efficiently manage the complexity of the exponential number of variables, we solve the continuous relaxation of Model 2 (master problem) via a CG scheme. The master problem is given by Equations (11a) and (11b), and 0 ≤ \theta_s ≤ 1 ∀ s ∈ S. The optimal solution is given as LBCG, a valid lower bound to Model 1.

The master problem is initially solved using S' ⊆ S, an initial subset of set S, and the dual values \( \pi^*_i \), associated to constraint (11b), are found. New variables (sequences) are generated one-by-one by finding sets \( s^* \subseteq S \) such that the dual constraint

\[ \sum_{i \in s^*} \pi^*_i \leq 1 \quad (12) \]

is violated, such that

\[ \sum_{i \in \mathcal{I} \subseteq s^*} \pi^*_i > 1 + \epsilon \]

Set \( s^* \) is found by solving Model 3, the subproblem, which finds a feasible set of transmission pairs of maximum violation. For the mathematical formulation, we define binary decision variables:

\[ \xi_i = \begin{cases} 
1, & \text{if pair } i \in \mathcal{I} \text{ is present in the set } s^* \\
0, & \text{otherwise}
\end{cases} \]

and decision variables \( \mu_i \in \mathbb{R}^+ \), the power level of pair i ∈ \mathcal{I}.

**Model 3 Subproblem**

\[
\begin{align*}
\text{maximise} & \quad \sum_{i \in \mathcal{I}} \pi^*_i \xi_i \quad (13a) \\
\text{subject to} & \quad \xi_i = 0 \Rightarrow \mu_i = 0 \quad \forall i \in \mathcal{I} \quad (13b) \\
& \quad \xi_i = 1 \Rightarrow g_{ii} \mu_i \geq \gamma_i \left( \sum_{j \in \mathcal{I}, j \neq i} g_{jj} \mu_j + v_i \right) \quad \forall i \in \mathcal{I} \quad (13c) \\
& \quad \xi_i \in \{0, 1\} \quad \forall i \in \mathcal{I} \quad (13d) \\
& \quad \mu_i \in \mathbb{R}^+ \quad \forall i \in \mathcal{I} \quad (13e)
\end{align*}
\]

The objective function (13a) aims at finding the maximum violation associated to a feasible set of transmission pairs. Constraints (13b) and (13c) ensure that if a transmission pair is not present in the set, then its associated power level is zero, and for all pairs in the set, the SINR conditions are satisfied, respectively. Finally, constraints (13d) and (13e) define the allowable values for the decision variables.

The master problem and subproblem are solved repeatedly until no more sequences violating constraint (12) can be found. Significant improvements in terms of the speed of the CG methodology are achieved by constructing and inserting a population of feasible sequences before the CG is invoked. The population includes all the sequences, which include up to two pairs, found using complete enumeration. We also include a subset of the sequences with at least three pairs, constructed using the heuristic approach described in Section 6.1.2.

Note that, by the end of the CG approach, a fractional value for the master problem may be found. In such a case, the value can be rounded up to the nearest integer, thus guaranteeing feasibility of the solution.

Finally, at the end of the CG approach, a number of feasible sets of transmission pairs is found, including the ones that were enumerated prior to the CG as well as the ones found using the subproblem (Model 3). Let \( \hat{S} \) be the collection of all these sets. If we solve Model 2 by substituting \( S \) with \( \hat{S} \subseteq S \), then the solution value found will be an upper bound to the original problem. This upper bound is denoted by UBCG.

### 6.1.4 Branch and bound approach

For the purposes of this paper, we implement B&B approaches, which at the initialisation stage set their lower bound as the maximum of lower bounds LB and LB', and the upper bound (cut-off value) as the minimum of upper bounds UB, UB' and UB''. It is possible also to add feasible cuts, that is, constraints that aim to exclude any solutions that we are certain to be non-optimal. For example, we may add constraints that force infeasibility...
on the solutions found with our upper bound techniques. As a result, the B&B algorithm will try to find the best solution not equal to the latter solutions, and if it cannot, then the excluded solutions were optimal. We classify such cuts as ‘pairwise’ and ‘all’, denoting feasibility cuts that disallow the simultaneous execution of pairs of nodes and of a number of nodes greater or equal to 2, respectively. For ‘pairwise’ cuts, we can find the pairs using complete enumeration (as we did in Section 6.1.1). For ‘all’ feasible cuts, we include the ‘pairwise’ cuts as well as the sets of solutions derived from the UB techniques. These are not strictly speaking ‘infeasible’, but we do not lose optimality by excluding them from the search because their solution value has been noted via the ‘cut-off’ value. We refer to the B&B approaches $BB_{paircuts}$ and $BB_{cuts}$, as the B&B algorithms incorporating ‘pairwise’ and ‘all’ cuts, respectively. For benchmarking purposes, we also use another B&B approach, referred to as BB, which is generic as it only implements CPLEX’s default settings (no bounds are incorporated).

### 7. PERFORMANCE EVALUATION

All the algorithms, as well as the LB and UB techniques described, have been implemented in Microsoft Visual Studio 2005 C++ using CPLEX v12.1 and run on an Intel Core 2 computer, with 2.5-GHz processor and 3.5 GB of RAM.

Throughout the paper, we set $y_1 = 3$ and $y_2 = 0.04$ mW. For each example, the selected number of nodes is uniformly and independently distributed on a square of side 100 m, making sure that no two nodes have distance between them less than 1 m. The links are constructed with a neighbor algorithm; that is, each transmitting node will choose a neighboring node as its receiver. Then, the channel gains $g_{ji}$ are obtained by considering distance attenuation only, that is, $g_{ji} = (d_0/d)^a$, where $d_0 = 1$ m and $a = 4$.

The algorithm begins with the computation of a lower bound (LB) and an upper bound (UB) according to the techniques described in Section 6.1.1 and the techniques described in Section 6.1.2, respectively. It is evident that if LB has the same value as UB, any optimisation technique is redundant; the heuristic solution would be optimal for the transmission scheduling problem. For the networks for which the calculated upper and lower bounds have different values (e.g. if upper bound = 5 and lower bound = 3), we use optimisation techniques in order to close the gap and, hence, obtain the optimal solution. At this stage, the initial pairwise cuts are found and added to the problem.

After determining the bounds, if the optimal solution is not found, we use B&B, which utilises the calculated bounds in order to derive the optimal solution faster. The apparent gain from determining as tight bounds as possible is the decrease in the CPU time required by the B&B technique. Apart from the optimisation techniques, a CG optimisation approach was designed in order to find stronger bounds for the transmission scheduling problem. The performance of CG was then compared with the other methods.

The performance of our algorithms was evaluated for 60 different networks of six different sizes: 10, 20, 30, 40, 50 and 60 pairs of nodes. For each size, we investigated 10 different networks. The performance of the algorithms is evaluated on three aspects: (a) the algorithms’ speed; (b) their scalability; and (c) their success at finding an optimal solution. Most bounding algorithms considered in the literature, consider only UB techniques, and therefore, a fair comparison can be made only for the UB methods proposed.

Figure 1 corresponds to the computational time for the LB and UB techniques for the networks considered of different sizes. The figure demonstrates that the computation time for both the lower and upper bounds is low for small networks and also it scales linearly with the number of communication pairs considered in the network. As a result, the algorithms succeed in having a low computational cost and good scalability properties.

![Figure 1. Average CPU times for the UB and LB algorithms. It is easy to see that the computational complexity of the algorithms scales linearly with the number of communication pairs in the network.](image-url)
In Figure 2, we compare the upper and lower bounds with the optimal value for each network considered, in order to get insight as to whether the non-trivial bounds designed perform well against the optimal values. For the figure, we can easily observe that for small networks, the bounding algorithms provide bounds that converge to the optimal value fast and with high probability, whereas for larger networks, the bounds are not tight but seem to remain close to the optimal value.

In Figure 3, it can be deduced that the UB algorithm produces slightly better results than the LB, because it matches the optimal solution more often than the LB algorithm, but they both remain close to the optimal value as the number of communication pairs increases. Our UB algorithm also outperforms $\text{ApproxLogN}$ [32].

For the networks for which the bounds algorithms (Sections 6.1.1 and 6.1.2) did not converge to the optimal value, we have used optimisation algorithms. From a total of 60 networks, the bounds algorithms (Sections 6.1.1 and 6.1.2) did not converge to the optimal value for 35 of them. It is remarkable to note that the lower bound, which was always the weakest bound, has comparable results to the upper bound; that illustrates that the derived non-trivial lower bound is tighter than existing results.

As shown in Table I, the computational cost for both lower and upper bounds is very small. When these bounds are incorporated into the B&B algorithm, we have significant improvements in the performance of the algorithms, and this will be shown next in Section 7.1.

### 7.1. Evaluation of the CG and B&B techniques

For the 35 networks whose optimal solutions were not found using the bounds, the CG method converged for nine networks (25.7%). Figure 4 compares the performance (in CPU time) of the CG and B&B approaches for these nine networks.

We can conclude that although CG is designed as an LB approach, in fact, all B&B variations are faster. The only exception are two networks, where CG performs slightly better than most of the B&B algorithms, except from the $BB_{\text{paircuts}}$. Given CG’s low optimality performance and its underperformance time-wise, we can infer...
that the scenarios where CG would be more appropriate than B&B or B&B with pairwise cuts are limited and negligible. Because of CG’s poor performance, we conclude that it is not an attractive approach for the solution of the scheduling problem.

8. DISCUSSION

In this paper, we studied the MLTS problem when power control can take place, that is, the problem of minimising the number of time slots required for scheduling all the wireless nodes in a given network for the abstract physical model when power control is allowed. The contributions of this paper are as follows.

1. We proved that this problem is NP-hard. Contrary to existing approaches and results, our formulation includes the choice of optimal transmitting powers and arbitrary topology. Furthermore, the generality of the abstract physical model implies the complexity of the geometric physical model and completes the theory that the MLTS problem with power control for the physical model in general is NP-hard. The authors of [30] and [35] prove NP-hardness for the geometric physical model without and with power control, respectively, and not the abstract physical model. For the abstract physical model, NP-hardness is proven in [36], but only for constant power levels. The NP-hardness proof for the abstract physical model with power control has been missing (the open problem posted in [38]). Through our proof, the conditions for which this problem becomes equivalent to graph colouring problem are presented. This would be helpful to know in scenarios in which we are asked to construct the network and hence, we will be able to avoid conditions that would make the computational complexity large.

2. We have developed non-trivial lower and upper bounds that can, in many instances, provide the optimal solution to the problem. In our simulations, it is illustrated that the upper bound is better in general than the lower bound and outperforms the current state-of-the-art approximation algorithm (ApproxLogN [32]). The lower bound was shown to be better than many other approaches (e.g. relaxations of optimisation formulations of the problem) but still remains an open problem.

3. Even if the bounds do not converge to the optimal solution, then when incorporated into the MILP formulation, it provides a considerable improvement in terms of computational time, as shown in the performance evaluation. As a result, the combined methodology scales to problems of non-trivial size. The CG approach (also appearing in many optimisation problems) has poor performance when compared with our problem-specific B&B approach, and hence, it is not an attractive approach for the solution of the scheduling problem. Other approaches, such as cutting plane, have been investigated (e.g. [1]), but they were proved inferior to B&B approaches.

9. CONCLUSIONS AND FUTURE DIRECTIONS

The MLTS problem in wireless networks for an intrinsically global model such as the abstract physical model with power control is NP-hard; it is thus unlikely to admit a polynomial-time optimal solution. The NP-hardness of the scheduling problem with power control for the abstract physical model, because of its generality, implies the NP-hardness for the geometric SINR model also. Therefore, the emphasis now is towards techniques that can provide strong, non-trivial lower and upper bounds for enhanced computational performance of analytical methods. To this end, we developed efficient bounding techniques that find good upper and lower bounds to the transmission scheduling problem. Further, we incorporated these bounds into a B&B implementation, showing that we are able to scale to problems of non-trivial size. Both the exact and heuristic approaches are useful in deriving the optimal solution value quickly, as well as providing feasible solutions of known quality in case the optimal solution is unknown. The significance of these results is threefold. On the one hand, the problem of transmission scheduling, where transmitters are able to adjust their power levels to fully benefit from spatial reuse, has been formulated and solved more efficiently with the aid of effective bounding techniques. On the other hand, the results are of practical importance in the presence of a central controller; the controller is able to make the calculations and disseminate the information to the rest of the network (e.g. in cellular networks where the base station can act as a centralised agent). Finally, the solution constitutes
an important benchmark when evaluating approximation algorithms or distributed algorithms for scheduling when knowledge of the whole network is unavailable.

Current research focuses on finding bounds that guarantee that the error lies within a maximum distance from the optimum solution or even prove hardness-of-approximation for an arbitrary gain matrix, that is, that no reasonable approximation algorithms can be developed for this problem. In addition, future work includes the implementation of the suggested algorithms on a field-programmable gate array whose computational performance will be compared with those coded on a PC using (IBM ILOG CPLEX Optimization Studio) CPLEX. Furthermore, the development of approximation algorithms for the transmission scheduling problem for the abstract physical model is part of ongoing research. Finally, a distributed algorithm poses a very challenging task and still remains an open problem.

ACKNOWLEDGEMENT

This research was sponsored in part by the Swedish Foundation for Strategic Research, SSF, under the RAM-COORAN project.

REFERENCES


