Reduced power expenditure in the minimum latency transmission scheduling problem

Kyriakos Deliparaschos†, Themistoklis Charalambous§, Paul Christodoulides† and Evelina Klerides#

†Electrical and Computer Engineering Department and Informatics, Cyprus University of Technology
§Department of Automatic Control, School of Electrical Engineering, KTH
#Business School, Imperial College London
Outline of the presentation

- Introduction
- Problem statement
- Problem formulation
- Numerical examples
- Conclusions and future directions
Medium Access Control (MAC)

MAC: Part of the Data Link Layer
- **Purpose**: to manage access to the shared wireless medium. Responsible for resolving conflicts among different nodes for channel access.
- **Nodes**: decide when to access the channel, avoiding collisions and efficiently utilizing the bandwidth.

Classification of MAC protocols:
- **Contention Based**
  - Random Access
  - Reservation/Collision resolution
- **Contention Free**
  - i.e. Transmission Scheduling
    (e.g. Token based, FDMA, CDMA, TDMA)
Transmission Scheduling

What is transmission scheduling?
- Time is divided into frames and frames into time-slots.
- Simplest example is TDMA (as many slots as nodes, one node per slot).

Why transmission scheduling?
Orchestrate channel access in order to fully exploit spatial reuse:
- Establish feasible networks for successful transmissions
- Minimize the number of time-slots required
- Minimize the power dissipated in the network
Problem Statement:
Given N communication requests, assign a color (time-slot) to each request. For all requests sharing the same color specify the power levels such that each request can be handled correctly (based on the abstract interference model).
We consider the *abstract* Physical Model where receivers experience *interference*:

\[ I_i = \sum_{j \neq i, j \in T} g_{ji}p_j + \nu \]

where

- \( g_{ij} \) is the channel gain on the link between transmitter \( i \) and receiver \( j \)
- \( p_i \) is the power level chosen by transmitter
- \( \nu \) is the variance of thermal noise at the receiver

The link quality is measured by the Signal-to-Interference-and-Noise-Ratio (SINR), given by

\[ \Gamma_i = \frac{g_{ii}p_i}{\sum_{j \neq i, j \in T} g_{ji}p_j + \nu} . \]

A transmission is successful (error free), if the SINR at the receiver is greater than the capture ratio, \( \gamma_i \):

\[ \frac{g_{ii}p_i}{\sum_{j \neq i, j \in T} g_{ji}p_j + \nu} \geq \gamma_i . \]
Assumptions:
- No node mobility or routing is considered.
- Time is slotted and each communication pair needs exactly one slot.
- Nodes maintain global synchronization: They know when slots and frames start.

Approach:
Interested in not only obtaining a transmission schedule of minimum duration, but also of minimum total power.

Decision variables:
- Processing-time variables:
  \[ x_i(t) = \begin{cases} 
  1, & \text{if transmitter } i \text{ is processed at time } t \\
  0, & \text{otherwise} 
\end{cases} \]
- Power level variables:
  \[ p_i(t) \in \mathbb{R}^+ \]
Formulation (1/2)

\[
\begin{align*}
\text{minimize} \quad & \tau = \max_{i \in \mathcal{T}} \sum_{t=1}^{D} tx_i(t) \\
\text{subject to} \quad & \sum_{t=1}^{D} x_i(t) \geq 1 \quad \forall i \in \mathcal{T}, \\
& x_i(t) = 0 \Rightarrow p_i(t) = 0 \quad \forall i \in \mathcal{T}, \ t = 1, \ldots, D, \\
& x_i(t) = 1 \Rightarrow g_{ii}p_i(t) \geq \gamma_i \left( \sum_{j \in \mathcal{T}, j \neq i} g_{ji}p_j(t) + \nu_i \right) \\
& \forall i \in \mathcal{T}, \ t = 1, \ldots, D, \\
& x_i(t) \in \{0, 1\} \quad \forall i \in \mathcal{T}, \ t = 1, \ldots, D, \\
& p_i(t) \in \mathbb{R}_+ \quad \forall i \in \mathcal{T}, \ t = 1, \ldots, D.
\end{align*}
\]

- $\tau$ is the latest time in time needed to schedule a transmission
- **1st constraint:** each link is scheduled, at least, once
- **2nd constraint:** if a link is not scheduled, the transmitting power is 0
- **3rd constraint:** if a link is scheduled, SINR constraint is fulfilled
- **4th-5th constraints:** define the admissible values of the decision variables
Formulation (2/2)

- Basic idea: Repeatedly solve a relaxation of the original problem, adding feasibility cuts.
- We partition the problem into a master problem and a subproblem.

\[
\begin{align*}
\text{minimize} & \quad \tau = \max_{i \in T} \sum_{t=1}^{D} tx_i(t) \\
\text{subject to} & \quad \sum_{t=1}^{D} x_i(t) \geq 1 \quad \forall i \in T \\
& \quad x_i(t) \in \{0, 1\} \quad \forall i \in T, \ t = 1, \ldots, D.
\end{align*}
\]

Master problem:
The iterative procedure solves the master problem - obtained optimal decision values to solve the subproblem.

\[
\begin{align*}
\text{minimize} & \quad \sum_{i \in T} \sum_{t=1}^{D} p_i(t) \\
\text{subject to} & \quad x_i^*(t) = 0 \Rightarrow p_i(t) = 0 \quad \forall i \in T, \ t = 1, \ldots, D, \\
& \quad x_i^*(t) = 1 \Rightarrow p_i(t)g_{ii} \geq \gamma_i \left( \sum_{j \in T, j \neq i} g_{ji}p_j(t) + \nu_i \right) \\
& \quad \forall i \in T, \ t = 1, \ldots, D, \\
& \quad p_i(t) \geq 0 \quad \forall i \in T, \ t = 1, \ldots, D.
\end{align*}
\]

Subproblem:
- If the subproblem is infeasible, new cuts are applied to the master problem.
- If the subproblem is feasible, optimal solution is guaranteed.
- Objective is changed, a good starting solution.
Proposed algorithm

Algorithm 1 A Cutting plane approach for the “Minimum latency transmission scheduling with SINR constraints” problem

initialise

\[ k = 1 \]

Compute \( LB \) \{Lower bound, e.g., from [8]\}
Compute \( UB \) \{Upper bound, e.g., from [8]\}
\( feas = false \) \{feasibility indicator, obtained from solving the Subproblem (Model 3)\}
Add cuts to Model 2.

if \( LB < UB \) then
repeat
Solve Model 2 to get optimal values \( x^*(k) \) and \( MP^*(k) \). Set \( LB = \max[MP^*(k), LB] \).
Solve Model 3 to check feasibility of SINR constraints.
if Model 3 is feasible then
\( feas = true \)
else
Add cuts of the form (9) to Model 2.
end if
\[ k = k + 1 \]
until \( feas = true \) or \( LB = UB \)
end if
\[ k = k - 1 \]
return \( LB, x^*(K) \).

Use heuristics to compute lower and upper bounds.
Solve Subproblem and check feasibility. If it is not feasible, add cuts to the Master problem.

The Master problem also provides a lower bound. So, the new LB is the maximum of what we found earlier and the solution to the Master problem.

If Subproblem finds a feasible solution, then the solution to the Master problem is optimal.

If Subproblem does not find a feasible solution, add cuts to the Master problem.

Examples - Numerical evaluations

Using $BB_{\text{slots}}$ (introduced in [8])

<table>
<thead>
<tr>
<th>Time Slot</th>
<th>Links in process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5 (15820.5), 8 (385019), 9 (579226)</td>
</tr>
<tr>
<td>2</td>
<td>3 (17739.8), 4 (763932), 7 (355380)</td>
</tr>
<tr>
<td>3</td>
<td>2 (206559), 6 (24638.4)</td>
</tr>
<tr>
<td>4</td>
<td>1 (2924.78), 10 (166.464)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of pairs</th>
<th>CPU time (sec) $CP_{\text{cuts}}$ $CP_{\text{slots}}$</th>
<th>$%$ out of time $CP_{\text{cuts}}$ $CP_{\text{slots}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.69 0.59</td>
<td>0 0</td>
</tr>
<tr>
<td>20</td>
<td>44.70 13.96</td>
<td>0 0</td>
</tr>
<tr>
<td>30</td>
<td>1840.18 587.23</td>
<td>3 2</td>
</tr>
<tr>
<td>40</td>
<td>N/A 548.20</td>
<td>10 4</td>
</tr>
</tbody>
</table>

**TABLE I**

Numerical Results for finding the optimal solution to Model 1 using $CP_{\text{cuts}}$ and $CP_{\text{slots}}$.

Using $CP_{\text{slots}}$

<table>
<thead>
<tr>
<th>Time Slot</th>
<th>Links in process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 (2924.02)</td>
</tr>
<tr>
<td>2</td>
<td>3 (2052.14), 6 (46050.3), 9 (276101)</td>
</tr>
<tr>
<td>3</td>
<td>4 (790562), 5 (64204.5), 7 (375443), 10 (4685.72)</td>
</tr>
<tr>
<td>4</td>
<td>2 (215659), 8 (122082)</td>
</tr>
</tbody>
</table>

Total power transmitted reduced by 20%.

$CP_{\text{cuts}}$ similar to $CP_{\text{slots}}$ but upper and lower bounding techniques (as in [8]) are not used.

$CP_{\text{slots}}$ but saves a lot of power, especially when the network is large, even though it does not guarantee to find the optimal configuration.

<table>
<thead>
<tr>
<th>No. of pairs</th>
<th>$%$ $\frac{power_{bb} - power_{cp}}{power_{cp}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>53%</td>
</tr>
<tr>
<td>20</td>
<td>3%</td>
</tr>
<tr>
<td>30</td>
<td>1646%</td>
</tr>
<tr>
<td>40</td>
<td>2409%</td>
</tr>
</tbody>
</table>

**TABLE II**

Numerical Results of total power required by the solution found by $BB_{\text{slots}}$ as compared to the solution found by $CP_{\text{slots}}$.

Conclusions and future directions

- A Cutting Plane approach is proposed:
  - Combines the minimum latency transmission scheduling problem
  - Partitions the problem into a Master Problem and a Subproblem
  - Aim to minimize the total power expenditure

- Methodology is bounded for large networks due to very large computation time
  - Parallel computation in GPGPU with OpenCL/CUDA

Significance of results:
- Transmitters able to adjust their power levels to fully benefit from spatial reuse. Of practical importance when a central controller exists.
- Important benchmark when evaluating heuristic or distributed algorithms for scheduling without knowledge of the whole network.
Thank you!

Questions?

For more information:
- k.deliparaschos@cut.ac.cy
- themisc@kth.se
- paul.christodoulides@cut.ac.cy
- e.klerides@imperial.ac.uk