Utility Based Resource Allocation in LTE Femtocell Cluster Considering Transport Network Limitations

27. ComNets-Workshop Mobil- und Telekommunikation und Treffen des Fördervereins der ComNets-Einrichtungen
13. March 2015
Outlines

• Introduction
• Resource allocation for non-real-time services
  – With transport limitations
    • aGW traffic shaping, problem solved by Lagrangian relaxation
    • Radio resource allocation, by utility based heuristics
  – No transport limitations: utility based optimal algorithm
• Simulation scenarios and results
• Summary and conclusion
Introduction

Figure: LTE architecture

EPC (Evolved Packet Core)

MME
S-GW/P-GW

MME
S-GW/P-GW

S1

eNB: E-UTRAN Node B
MME: Mobility Management Entity
S-GW: Serving Gateway
P-GW: PDN (Packet Data Network) Gateway
S1: The interface between eNBs and EPC
Introduction

Figure: Example of radio resource allocation

Figure: Example of aGW traffic shaping

- aGW (for S1 DL traffic)
  - IP level per-bearer shaping
Utility-based resource allocation

• Resource Allocation (RA): the amount of resources allocated to different users
  – Make the best use of limited resources under time varying channel conditions
  – Fairness, latency reduction, spectral efficiency and system utilization

• Utility-based Resource Allocation
  – Utility reflects actual users’ perceived performance (QoE)
  – Optimization problem: Maximize the aggregated utility, subject to limited resources
QoE examples

Figure: MOS over data rate [1][3]
## Utility functions

<table>
<thead>
<tr>
<th>Applications</th>
<th>Elastic traffics; Video with transcoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Properties</td>
<td>QoE monotonically increases with the data rate; Marginal QoE monotonically decreases with the data rate</td>
</tr>
<tr>
<td>Function Type</td>
<td>Sigmoid Function (Concave part)</td>
</tr>
<tr>
<td>Utility functions</td>
<td></td>
</tr>
</tbody>
</table>

\[ u(r) = \frac{A}{1 + e^{-\alpha r}} + B \]

![Utility curves](image)

*with A=9, B=-4.5*
General overview

Goal: \( \text{max} \left\{ U = \sum_c \sum_i u_{i,c} \left( b_{i,c} \right) \right\} \)

Transport network rate \( R_{S1} \)

\[ \sum_c \sum_i r_{i,c} = \sum_c \sum_i b_{i,c} \cdot \sigma_{i,c} \leq R_{S1} \]

Cell bandwidth \( B_c \)

\[ \sum_i b_{i,c} \leq B_c \quad \forall c \]

<table>
<thead>
<tr>
<th>Case</th>
<th>aGW traffic shaping</th>
<th>Radio scheduler</th>
</tr>
</thead>
<tbody>
<tr>
<td>No S1 bottleneck</td>
<td>-</td>
<td>Optimal algorithm</td>
</tr>
<tr>
<td>Only S1 bottleneck</td>
<td>Lagrangian relaxation solved by bisection search</td>
<td>Two heuristics (Centralized/Coordinated MAC scheduler)</td>
</tr>
<tr>
<td>Both S1 and some cells are bottleneck</td>
<td>Lagrangian relaxation solved by projected subgradient method</td>
<td></td>
</tr>
</tbody>
</table>
Lagrangian relaxation

- Maximize the aggregated utility in the cell cluster, which can be expressed as:

\[
\max \left\{ U = \sum_c \sum_i u_{i,c} \left( b_{i,c} \right) \right\}
\]

s.t. \( \sum_i b_{i,c} \leq B_c \quad \forall c; \quad \sum_c \sum_i b_{i,c} \cdot \sigma_{i,c} \leq R_{S1} \)

- It can be solved optimally using the Lagrangian decomposition method.
  - Hessian matrix positive definite \( \Rightarrow \) Problem is convex
  - Slater’s condition fulfilled \( \Rightarrow \) Strong duality holds

\[
f = \min_{\lambda} \left\{ \max_{b} \left\{ \sum_c \sum_i u_{i,c} \left( b_{i,c} \right) - \sum_c \lambda_c \left( \sum_i b_{i,c} - B_c \right) - \lambda_0 \left( \sum_c \sum_i b_{i,c} \cdot \sigma_{i,c} - R_{S1} \right) \right\} \right\}
\]

\[
= \min_{\lambda} \left\{ \sum_c \sum_i \max_{b} \left\{ u_{i,c} \left( b_{i,c} \right) - \left( \lambda_c + \lambda_0 \cdot \sigma_{i,c} \right) b_{i,c} \right\} + \sum_c \lambda_c \cdot B_c + \lambda_0 \cdot R_{S1} \right\}
\]

\[
= \min_{\lambda} \left\{ \sum_c \sum_i L^*_{i,c} + \sum_c \lambda_c \cdot B_c + \lambda_0 \cdot R_{S1} \right\}
\]

- Subgradient projection method is applied
  - with modified Polyak’s step size
Visualization of the subgradient method

\[ \text{minimize } f(\lambda) \]

\[ \text{subject to } \lambda \in \mathbb{R}^+ \cup \{0\} \]

\[ f = \sum_c \sum_i L^*_i,c + \sum_c \lambda_c \cdot B_c + \lambda_0 \cdot B_{S1} \]

Figure: Visualization of the subgradient method

Figure: Visualization of convergency
Radio scheduler heuristics

Figure: The algorithm of the proposed heuristic

Heuristic Algorithm 1:
\[ m_i = \Delta u_{i,n+1} - \Delta u_{i,n} \]

Heuristic Algorithm 2:
\[ m_i = \frac{\Delta u_{i,n+1} - \Delta u_{i,n}}{\sigma_i} \]

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i,c)</td>
<td>User index; Cell index</td>
</tr>
<tr>
<td>(M:C)</td>
<td>Number of users; Number of cells</td>
</tr>
<tr>
<td>(N_i)</td>
<td>Number of PRBs allocated to user (i)</td>
</tr>
<tr>
<td>(N_{a,c})</td>
<td>Total available PRBs in the cell (c) after HARQ (Hybrid Automatic Repeat Request)</td>
</tr>
<tr>
<td>(N_{cell})</td>
<td>Number of PRBs in the cell</td>
</tr>
<tr>
<td>(m_i)</td>
<td>QoE metric of user (i)</td>
</tr>
<tr>
<td>(u_i)</td>
<td>QoE of user (i)</td>
</tr>
<tr>
<td>(\Delta u_{i,n})</td>
<td>Marginal QoE of user (i) with (n) PRBs</td>
</tr>
<tr>
<td>(\sigma_i)</td>
<td>Channel quality indicator of user (i)</td>
</tr>
<tr>
<td>(R_S1)</td>
<td>Transport network capacity in Mbps</td>
</tr>
<tr>
<td>(R_{S1,e})</td>
<td>Estimated achievable throughput on (S_1)</td>
</tr>
</tbody>
</table>
Simulation tool and scenario

Use OPNET discrete event simulation software (v17.5)
Includes modeling of
- E-UTRAN and EPC entities
- Full protocol stack including MAC, RLC, PDCP, GTP etc.
- Mobility models
- IP DiffServ
## Simulation settings

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Settings</th>
</tr>
</thead>
</table>
| **Macro eNBs (cells) settings** | 7 eNBs with hexagonal coverage, 500ms inter-eNB distance (center eNB located at the original point (0m, 0m))  
Path loss: 130.5 + 37.6log10(R), R in Km [5]  
Slow fading: Correlated Log normal, zero mean, 8db std. and 50 m correlation distance  
Small scale fading: 3GPP Pedestrian A  
Transmission power: 23dBm per PRB                                                                                     |
| **Femtocell cluster settings** | Building size with 40mx40m, center coordinate: (200m,0m)  
3 femtocell station coordinates:  
(210m, -10m), (190m, 10m), (210m,10m)  
Penetration loss (interference from macro eNBs) over the wall: 12dB mean with 8dB std.  
Path loss: 41.1 + 16.9*log10(R), R in Km [5]  
Small scale fading: 3GPP Pedestrian A  
Transmission power: 0dBm per PRB                                                                                      |
| **TCP version**                | New Reno with 64Kbytes receiver buffer size                                                                                                                                                           |
| **Traffic types**              | VoIP: GSM EFR, codec rate 12.2 kbps  
Video Streaming: TCP based full buffer streaming  
HTTP: 2MB page size, Inter arrival time: exp. distributed with mean: 50s  
FTP: 10MB file size, Inter arrival time: exp. distributed with mean: 50s                                                                 |
| **Mobility model**             | 5Km/h, Random waypoint                                                                                                                                                                                  |
| **aGW shaper**                 | Token Bucket algorithm, maximum token bucket size: 64KB                                                                                                                                             |
| **Transport limitation**       | 16Mbps or 6 Mbps (aGW->femtocell cluster); 1Mbytes buffer size                                                                                                                                         |
| **Number of PRBs**             | 25 PRBs (5MHz spectrum at 2.6 GHz)                                                                                                                                                                     |
| **Simulation time**            | 1000s (5 runs with different seeds) with warm up period of 300s                                                                                                                                        |
QoE functions

\[ u = \frac{A}{1 + e^{-\alpha r}} + B \]

<table>
<thead>
<tr>
<th>Activity</th>
<th>A</th>
<th>B</th>
<th>( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video</td>
<td>6.954</td>
<td>-2.542</td>
<td>4.104</td>
</tr>
<tr>
<td>Web</td>
<td>5.815</td>
<td>-1.735</td>
<td>1.294</td>
</tr>
<tr>
<td>FTP</td>
<td>8</td>
<td>-3</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table: Parameters for QoE Functions

Fig. QoE functions used in the simulation [1][3]
Scenario 1 - low load

- Voice packet end-to-end delay (msec): 1, 2, 3, 4
- Video streaming data rate (Mbps): 3, 6, 8, 12
- HTTP page response time (sec): 8, 16, 24, 32
- FTP file upload response time (sec): 16 Mbps

PF: Propotional Fair
Scenario 1 - low load

The mean opinion score (MOS) for different methods under low load conditions is shown in the graph. The methods compared include:

- Video Streaming
- HTTP
- FTP
- Total

The graph indicates that the total mean opinion score (MOS) is highest for the aGW shaping method, followed by MAC Heuristic 2, MAC Heuristic 1, and PF respectively.
Scenario 2 – high load

- Video streaming data rate (Mbps): 8, 16, 24, 32
- HTTP page response time (sec): 48, 96, 128, 192
- FTP file upload response time (sec): 48, 96, 128, 192

- Voice packet end-to-end delay (msec): 0.25, 0.5, 0.75, 1
- Video streaming data rate (Mbps): 0.25, 0.5, 0.75, 1

- Network configurations:
  - 10 HTTP
  - 10 FTP
  - 10 VoIP
  - 10 Video Streaming

- Bitrate: 16 Mbps
Scenario 2 – high load

![Graph showing Mean Opinion Score (MOS) for different methods with error bars for Video, HTTP, FTP, and Total Mean Opinion Score (MOS).]
Scenario 3 – very high load

Mean Opinion Score (MOS)

Method

PF  MAC Heuristic 1  MAC Heuristic 2  aGW shaping

Total Mean Opinion Score (MOS)

10 HTTP  10 FTP  10 VoIP  10 Video Streaming

6 Mbps

Video Streaming
HTTP
FTP
Total

29.04.2015
Summary

- A utility based resource allocation framework for LTE is proposed.
- The problem is formulated as a convex optimization problem and analytically solved using Lagrangian decomposition method.
- The approach is implemented in the simulator. The performance of utility based approach is compared against PF scheduler.

<table>
<thead>
<tr>
<th>Case</th>
<th>aGW traffic shaping</th>
<th>Radio scheduler</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both S1 and some cells are bottleneck</td>
<td>Lagrangian relaxation Solved by projected subgradient method</td>
<td>Two heuristics (Centralized/Coordinated MAC scheduler)</td>
</tr>
<tr>
<td>Advantages</td>
<td>1. Give the best performance 2. No need to modify the radio scheduler</td>
<td>1. Low computational power 2. Good performance, heuristic 2 is better than 1 in high load scenarios</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>1. Signallings between eNB and aGW 2. High computational power</td>
<td>1. Need a centralized/coordinated scheduler among the cells sharing the same transport link</td>
</tr>
</tbody>
</table>
Ongoing works

\[
\max \left\{ U = \sum_c \sum_i w_{i,c} \cdot u_{i,c} \left( r_{i,c} \right) \right\}
\]

<table>
<thead>
<tr>
<th>Traffic Type</th>
<th>Delay sensitive traffics (Real-time)</th>
<th>Rate sensitive traffics (Non Real-time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility Functions</td>
<td>( u_{i,c} \left( r_{i,c} \right) = \frac{u'(d_{i,c})}{\lambda_{i,c}} \cdot r_{i,c} )</td>
<td>( u_{i,c} \left( r_{i,c} \right) = \frac{A}{1 + e^{-\alpha_{i,c} r_{i,c}}} \cdot r_{i,c} + B )</td>
</tr>
<tr>
<td>Optimization Model</td>
<td>Linear Programming</td>
<td>Concave Optimization</td>
</tr>
</tbody>
</table>

![Graph](image)
References

1. ITU-T recommendation G.1070: Opinion model for videotelephony applications (04/2007)
Thanks and Any Questions?