Switching Constrained Max-Weight Scheduling for Wireless Networks
Soumya Basu, and Sanjay Shakkottai

The University of Texas at Austin
Electrical and Computer Engineering
 Cockrell School of Engineering

Dense Cellular Networks
• Dense deployment of base stations (BS) to support peak data traffic
• Dynamic activation and de-activation of BS to optimize energy usage
• Fast activation dynamics is necessary to maintain data rate
• Fast activation dynamics leads to large switching overhead, e.g. hand-offs, state exchange among BSs, and BS start-up costs

Joint Channel Realizations

<table>
<thead>
<tr>
<th>Probability μ_h: 0.5</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Arrival Queue</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BS Activation</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Scheduling Queue</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Departure Queue</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Need for a New Approach
• Greedy Optimization Techniques (Primal-Dual, Drift+Penalty)
• Frequent BS state change as switching cost is not optimized
• Reinforcement learning with bounded queue length
• Prohibitive computation for large queue lengths
• Complex packet drop vs optimality tradeoffs
• Sticky BS selection using static-split rule + MW scheduling
• Large delay as BS selection is non-adaptive to queue lengths

Learning Aided Switching and Scheduling (LASS)
• Parameters: Switching rate, ε_s and Penalty scale, V
• Independent R.Vs: Switch: E_{t}(t)={Ber(ε_s)}, Explore: E_{t}(t)={Ber(log(t)/t)}

BS Activation and Scheduling
• Two key decisions: BS activation and Channel Scheduling
• BS activation is further split into two decisions
  • When to switch? Switch at a very slow rate (Constrained Switching)
  • What to switch to? Expected Max-weight with activation penalty
• Channel state is unknown before BS activation
• Exploration-exploitation tradeoff in learning channel states
• Channel Scheduling: Exact max-weight with known channel state

Dynamic Switching
• Uses two variables Switch Counter: T(t) and Switch Queue : Q_{sw}(t)
• Switch counter keeps count of the time since \( f(t) \) is scheduled
• Switch queue counts the number of switching events that exceeds rate \( ε_s \)

\[
\text{Switch queue: } Q_{sw}(t+1) = (Q_{sw}(t) - E_{t}(t) + 1)\{Q_{sw}(t) \leq T(t)\}
\]

\[
\text{Switch counter: } T(t+1) = 1 + (f(t) - f(t-1) + 1)\{T(t) \leq f(t)\}
\]

Performance Guarantees
• Assumptions on system parameters
  • Capacity gap \( ε_s > 0 \), and bounded arrivals and departures
  • Optimal cost of the system \( C_{avg} \) with no switching constraints
• Algorithm parameters: Switching rate, \( ε_s \) and Penalty scale, V
• Performance metrics of interest:
  • Time average of queue lengths: \( Q_{avg} \) and costs: \( C_{avg} \)
  • Tail bounds for queue lengths: \( P(\{Q(t) \leq \tilde{x}\}) \)

Simulation Results
• Four algorithms are simulated for 8 Users and 3 BSs until convergence:
  • DP: Greedy Drift + Penalty
  • LSG: LASS Static with geometric inter arrival between \( E_{t}(t) \)
  • LSF: LASS Static with fixed inter arrival between \( E_{t}(t) \)
  • LD: LASS with Dynamic switching
• First plot: \( Q_{avg} \) of DP < LD < LSF < LSG (V = 100, load = 0.9)
• Second plot: \( C_{avg} \) of DP > LD \approx LSF \approx LSG (V = 100, \( ε_s = 0.1 \))
• Third and Fourth plot: Separation of queue length tail distribution
  • DP < LD < LSF < LSG (V = 100, load = 0.9)
• Differences are more pronounced for smaller \( ε_s \)