SLOW RESOURCE ALLOCATION FOR HETEROGENEOUS NETWORKS

Michael Honig
Department of EECS
Northwestern University

Joint work with Binnan Zhuang, Dongning Guo
Mobile Data Forecast


CUHK, April 2016
Evolution to 5G

- New spectrum (mm-wave, unlicensed)
- Physical layer advances
  - massive MIMO, network coding, cooperation
- Smaller cells
Heterogeneous Network

User Deployed WiFi Access Points/Femtocells/Relays

Operator Deployed Pico cells/Relays

Remote Radio heads
Remote Radio Heads

How to do interference management?
Offline Frequency Planning (1G-4G)
Slow Resource Allocation

- Over many packets (seconds)
  - Average channel gains, offered traffic
- Combined with fast scheduling (milliseconds)
- Traffic varies over space, stationary in time
- Centralized approach

**Contribution:** general optimization framework
Cells overlap, traffic varies.
How to allocate spectrum across cells?
Assumptions

- Resources within each cell are allocated via fast scheduling.
- Resources across cells are allocated over a slower time-scale.
- Centralized controller knows average traffic, average channels.
Consider all possible ways the spectrum can be partitioned among BTS’s.

Optimize over this partition.
Two Base Stations

Traffic for users in cell 1

\[ \lambda_1 \rightarrow \]

Traffic for users in cell 2

\[ \lambda_2 \leftarrow\]

BTS 1

\[ \leftarrow x\{1\} \rightarrow x\{1,2\} \rightarrow x\{2\} \rightarrow \]

BTS 2

BW assigned to BTS 1

| BW assigned to both BTS 1 and 2 | BW assigned to BTS 2 |

Total available bandwidth (BW)

CUHK, April 2016
Orthogonal Allocation

Traffic for users in cell 1

\[ \lambda_1 \rightarrow \]

Traffic for users in cell 2

\[ \lambda_2 \leftarrow \]

BTS 1

\[ x\{1\} \]

BTS 2

\[ x\{2\} \]

BW assigned to BTS 1

BW assigned to BTS 2

Total available bandwidth (BW)

CUHK, April 2016
Full Frequency Reuse

Traffic for users in cell 1

\( \lambda_1 \rightarrow \)

Traffic for users in cell 2

\( \lambda_2 \leftarrow \)

BTS 1

BTS 2

All BW assigned to both BTS 1 and 2

Total available bandwidth (BW)

CUHK, April 2016
Partial Sharing

Traffic for users in cell 1

\[ \lambda_1 \rightarrow \text{BW assigned to BTS 1} \]

Traffic for users in cell 2

\[ \lambda_2 \leftarrow \text{BW assigned to BTS 2} \]

\[ x\{1\} \quad x\{1,2\} \quad x\{2\} \]

Total available bandwidth (BW)

CUHK, April 2016
Orthogonal Allocation

Traffic for users in cell 1

\(\lambda_1\)

\(x\{1\}\)

BW assigned to BTS 1

Traffic for users in cell 2

\(\lambda_2\)

\(x\{2\}\)

BW assigned to BTS 2

"I would build a GREAT wall!"

Total available bandwidth (BW)
Full Frequency Reuse

Traffic for users in cell 1

Traffic for users in cell 2

\( \lambda_1 \rightarrow \)  

\( \lambda_2 \leftarrow \)

"Tear down this wall!"

\( x \{1, 2\} \)

All BW assigned to both BTS 1 and 2

Total available bandwidth (BW)

CUHK, April 2016
Partial Sharing

Traffic for users in cell 1

\[ \lambda_1 \rightarrow \]

\[ x_1 \]

Traffic for users in cell 2

\[ x_{1,2} \]

\[ \rightarrow \lambda_2 \]

"one country, two systems"

BW assigned to BTS 1

Shared BW

BW assigned to BTS 2

Total available bandwidth (BW)

CUHK, April 2016
Partial Sharing

Traffic for users in cell 1

$\lambda_1 \rightarrow \mathcal{X}\{1\}$

Traffic for users in cell 2

$\lambda_2 \leftarrow \mathcal{X}\{2\}$

$\mathcal{X}\{1,2\}$

“one country, two systems”

<table>
<thead>
<tr>
<th>BW assigned to BTS 1</th>
<th>Shared BW</th>
<th>BW assigned to BTS 2</th>
</tr>
</thead>
</table>

Partition should depend on traffic!
3-BTS Example
K Base Stations

spectrum allocation: $2^k$ reuse patterns (variables)

$\lambda_1 \rightarrow \Lambda_1 \rightarrow \Lambda_2 \rightarrow \Lambda_3$
Bandwidth Optimization Problem

- Adjust partition to minimize average latency
- Take into account queuing delays and interference
- Interference affects achievable rates

Rate per Hz: $S_{1,\{1\}}$, $S_{1,\{1,2\}} \leq S_{1,\{1\}}$
Bandwidth Optimization Problem

- Adjust partition to minimize average latency
- Take into account queuing delays and interference
- Interference affects achievable rates

Rate per Hz: $S_{1,\{1\}}$  $S_{1,\{1,2\}}$  $S_{2,\{1,2\}}$  $S_{2,\{2\}}$

BW assigned to BTS 1  BW assigned to both BTS 1 and 2  BW assigned to BTS 2

CUHK, April 2016
Spectral Efficiency

$$s^{i \rightarrow j}_{A} = 1_{i \in A} \frac{W}{L} \log \left( 1 + \frac{p^{i \rightarrow j}_{IA \rightarrow j}}{I_{A \rightarrow j}} + \sigma^2 \right)$$

- Average powers, channels
- Known to the optimizer
Bandwidth Optimization Problem

BTS 1 transmits

BTS 2 transmits

Rate per Hz: $s_{1,\{1\}}, s_{1,\{1,2\}}, s_{2,\{1,2\}}, s_{2,\{2\}}$

BW assigned to BTS 1

BW assigned to both BTS 1 and 2

BW assigned to BTS 2

Total rates:

$r_1 = s_{1,\{1\}} x_{\{1\}} + s_{1,\{1,2\}} x_{\{1,2\}}$

$r_2 = s_{2,\{2\}} x_{\{2\}} + s_{2,\{1,2\}} x_{\{1,2\}}$
Bandwidth Optimization Problem

BTS 1 transmits

Rate per Hz: $S_1, \{1\}$

BW assigned to BTS 1

BTS 2 transmits

Rate per Hz: $S_1, \{1,2\}$

BW assigned to both BTS 1 and 2

Rate per Hz: $S_2, \{1,2\}$

BW assigned to BTS 2

Rate per Hz: $S_2, \{2\}$

Total rate from BTS $i$: $r_i = \sum_{B \subset N} S_{i,B} x_B$

sum over all reuse patterns

$\mathcal{N} = \{1, 2, \cdots, N\}$ set of BTSs

CUHK, April 2016
Bandwidth Optimization Problem

- Adjust partition to minimize average latency
- Take into account queuing delays and interference
- Interference affects achievable rates
- Queues at different BTS’s are dependent – complicates optimization!

BTS 1
sometimes
transmits

BTS 2
sometimes
transmits
Backlogged Traffic: Delay

BTS 1 transmits

BTS 2 transmits

Rate per Hz: $S_{1,\{1\}}$ $S_{1,\{1,2\}}$ $S_{2,\{1,2\}}$ $S_{2,\{2\}}$

- BW assigned to BTS 1
- BW assigned to both BTS 1 and 2
- BW assigned to BTS 2

Average packet sojourn time (M/M/1): $t_i = \frac{1}{r_i - \lambda_i}$
Conservative Optimization

\[
\min_{\{x,r\}} \sum_{i=1}^{N} \left( \frac{\lambda_i}{\sum_{i=1}^{N} \lambda_i} \right) \frac{1}{r_i - \lambda_i}
\]

Subject to:
\[
\begin{align*}
& r_i > \lambda_i \\
& r_i = \sum_{B \subseteq \mathcal{N}} s_{i,B}x_B \quad \forall i \in \mathcal{N} \\
& x_B \geq 0 \quad \forall B \subseteq \mathcal{N} \\
& \sum_{B \subseteq \mathcal{N}} x_B = 1
\end{align*}
\]

- Convex, \(2^N-1\) variables
- The solution achieves the maximum throughput region.
Property of Solution

- **Theorem:** The optimal allocation divides the spectrum into at most $N$ segments (instead of $2^N$).
- Follows from Carathéodory’s theorem.
- 7-BTS example:
Interactive Queues (Two BTSs)
Assumptions:

- The $N$ queues are independent conditioned on the pattern.
- For the transition $A \rightarrow A'$, the new state is chosen according to the steady-state distribution.
Refined Optimization

\[
\min_{\{x,r,t\}} \sum_{i=1}^{N} \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} t_i
\]

Subject to:

\[
t_i = \sum_{A \ni i} \frac{p(A)r_{i,A}}{\lambda_i(r_{i,A} - \lambda_i)}
\]

\[
r_{i,A} = \sum_{B \subset \mathcal{N}} s_{i,B \cap A x B}, \quad r_{i,N} > \lambda_i
\]

- Not convex
- The solution achieves the maximum throughput region.
One-Dimensional Example

Spectrum Allocation

Traffic load

CUHK, April 2016
Small-Cell Network

7 BTS’s
100x100 m²
Pkt. length 1MB
Bandwidth 160MHz
Pathloss exponent 3
Delay vs. Traffic Intensity

- **Orthogonal**
- **Full-Spectrum-Reuse**
- **Conservative**
- **Refined**

- 7 BTS’s
- 100x100 m²
- Pkt. length 1MB
- Bandwidth 160MHz
- Pathloss exponent 3

Average packet sojourn time (seconds) vs. Average packet arrival rate per BTS (packets/sec)

CUHK, April 2016
Approximation vs. Bounds

![Graph showing comparison between simulation and approximations](image)

- **Simulation**
- **Refined Approximation**
- **Second-degree Upper Bound**
- **Second-degree Lower Bound**
- **Conservative Approximation**

CUHK, April 2016
Conservative vs. Refined Allocations

CUHK, April 2016
Load to BTS Assignment

Problem: Jointly allocate traffic and bandwidth across base stations.

Optimization variables:
Spectrum used by BTS $n$ to serve hexagon $k$ under reuse pattern $A$
Load to BTS Assignment: Notation

- Set of BTSs: $\mathcal{N} = \{1, 2, \ldots, N\}$
- Set of UE groups: $\mathcal{K} = \{1, 2, \ldots, K\}$
- $\lambda_k$: packet arrival rate for group $k$
\( S_A^{i \rightarrow j} \): spectral efficiency of BTS \( i \) serving group \( j \) under reuse pattern \( A \).

\( x_A^{i \rightarrow j} \): spectrum resource used by BTS \( i \) to serve group \( j \) under reuse pattern \( A \).

\( y_A \): fraction of spectrum resources allocated to reuse pattern \( A \).
Conservative Optimization (Original)

\[
\min_{\{x,r\}} \sum_{i=1}^{N} \left( \frac{\lambda_i}{\sum_{i=1}^{N} \lambda_i} \right) \frac{1}{r_i - \lambda_i}
\]

Subject to:

- \( r_i > \lambda_i \)
- \( r_i = \sum_{B \subset \mathcal{N}} s_{i,B} x_B \quad \forall i \in \mathcal{N} \)
- \( x_B \geq 0 \quad \forall B \subset \mathcal{N} \)
- \( \sum_{B \subset \mathcal{N}} x_B = 1 \)
Conservative Optimization (Modified)

\[
\max_{\mathbf{x}, \mathbf{r}} U(\mathbf{x}, \mathbf{r})
\]

Subject to:
\[
\begin{align*}
  r_i &> \lambda_i \\
  r_i &= \sum_{B \subseteq \mathcal{N}} s_{i,B} x_B \quad \forall i \in \mathcal{N} \\
  x_B &\geq 0 \quad \forall B \subset \mathcal{N} \\
  \sum_{B \subseteq \mathcal{N}} x_B &= 1
\end{align*}
\]
Conservative Optimization (Modified)

\[
\max_{x, r} U(x, r)
\]

Subject to:
\[
\forall j \in \mathcal{K} : r^j = \sum_{i=1}^{n} \sum_{B \subset \mathcal{N}} s^i_{B \rightarrow j} x^i_{B \rightarrow j}
\]
\[
x_B \geq 0 \quad \forall B \subset \mathcal{N}
\]
\[
\sum_{B \subset \mathcal{N}} x_B = 1
\]

CUHK, April 2016
Conservative Optimization (Modified)

\[
\max_{\mathbf{x}, \mathbf{r}} U(\mathbf{x}, \mathbf{r})
\]

Subject to:

\[
\begin{align*}
    r^j &= \sum_{i=1}^{n} \sum_{B \subseteq \mathcal{N}} s_{B}^{i \rightarrow j} x_{B}^{i \rightarrow j} & \forall j \in \mathcal{K} \\
    \sum_{j=1}^{K} x_{B}^{i \rightarrow j} &= y_B & \forall i \in \mathcal{N} \\
    \sum_{i=1}^{\mathcal{N}} \mathbf{1} &\geq 0
\end{align*}
\]

- Convex for concave \( U, O(KN2^N) \) variables
- The solution achieves the maximum throughput region.
Average Delay Minimization

\[
\min \sum_{j=1}^{K} \lambda^j \frac{1}{(r^j - \lambda^j)^+}
\]

Subject to:

\[
r^j = \sum_{i=1}^{n} \sum_{B \subseteq \mathcal{N}} s_{i \rightarrow j}^B x_{i \rightarrow j}^B \quad \forall j \in \mathcal{K}
\]

\[
\sum_{j=1}^{K} x_{i \rightarrow j}^B = y_B \quad \forall i \in \mathcal{N}
\]

\[
\sum_{j=1}^{K} \lambda^j > 0
\]

- Convex for concave \( U, O(\mathcal{K}2^\mathcal{N}) \) variables
- The solution achieves the maximum throughput region.
Properties of the Solution

- Uses at most $K$ of the $2^N$ reuse patterns
- At most $N-1$ groups are jointly served by $\geq 1$ AP.
- Throughput optimal
Spectrum Allocation

CUHK, April 2016
Delay (2 macros, 8 small cells)

![Graph showing delay vs. average packet arrival rate per user type]

- **Full spectrum reuse + max RSRP**
- **Full spectrum reuse + intra-cell allocation**
- **Inter-cell allocation + max RSRP**
- **Inter-cell allocation + intra-cell allocation**

CUHK, April 2016
Energy Conservation

- **Tradeoff:**
  - Turning off an AP saves energy.
  - Turning off an AP increases the load for neighbors.

- **Problem:** serve the offered traffic with the minimal number of active APs.

- **Related work:**
  [Pollakis, Cavalcante and Stanczak, ’12]
  (no spectrum optimization)
Average Delay Minimization

$$\min \sum_{j=1}^{K} \lambda^j \frac{1}{(r^j - \lambda^j)^+}$$

Subject to:

$$r^j = \sum_{i=1}^{n} \sum_{B \subseteq \mathcal{N}} s_{i \rightarrow j}^{B} x_{i \rightarrow j}^{B} \quad \forall j \in \mathcal{K}$$

$$\sum_{j=1}^{K} x_{B \rightarrow j}^{i} = y_B$$

$$\sum_{B \subseteq \mathcal{N}} y_B = 1, \quad \mathbf{x} \geq 0$$
Weighted Energy Minimization

\[ \min_z \sum_{i \in \mathcal{N}} c^i |z^i|_0 \]

Subject to:

- total bandwidth assigned to AP \( i \)

\[ r^j = \sum_{i=1}^{n} \sum_{B \subset \mathcal{N}} s_{B}^{i \rightarrow j} x_{B}^{i \rightarrow j} \quad \forall j \in \mathcal{K} \]

\[ \sum_{j=1}^{K} x_{B}^{i \rightarrow j} = y_B, \quad \sum_{B \subset \mathcal{N}} y_B = 1 \]

\[ \mathbf{x} \geq 0 \]
Weighted Energy Minimization

\[
\min_{z} \sum_{i \in \mathcal{N}} c^i |z^i|_0
\]

Subject to:

\[
\sum_{B \subseteq \mathcal{N}} \sum_{j \in \mathcal{K}} x_{i \to j}^B \leq z^i \quad \forall i \in \mathcal{N}
\]

\[
r^j - \lambda^j \geq \frac{1}{\tau^j}
\]

\[
r^j = \sum_{i=1}^{n} \sum_{B \subseteq \mathcal{N}} s_{i \to j}^B x_{i \to j}^B \quad \forall j \in \mathcal{K}
\]

\[
\sum_{j=1}^{K} x_{i \to j}^B = y_B, \quad \sum_{B \subseteq \mathcal{N}} y_B = 1
\]

\[
x \geq 0
\]

CUHK, April 2016
Weighted Energy/Utility Minimization

\[ \min_{z, r} \sum_{i \in \mathcal{N}} c^i |z^i|_0 + U(r) \]

Subject to:

\[ \sum_{B \subseteq \mathcal{N}} \sum_{j \in \mathcal{K}} x_{B}^{i \rightarrow j} \leq z^i \quad \forall i \in \mathcal{N} \]

\[ r^j - \lambda^j \geq 1/\tau^j \]

\[ r^j = \sum_{i=1}^{n} \sum_{B \subseteq \mathcal{N}} s_{B}^{i \rightarrow j} x_{B}^{i \rightarrow j} \quad \forall j \in \mathcal{K} \]

\[ \sum_{j=1}^{K} x_{B}^{i \rightarrow j} = y_B, \quad \sum_{B \subseteq \mathcal{N}} y_B = 1 \]

\[ x \geq 0 \]
Reweighted $\ell_1$ Minimization

\[
\min_{z,r} \sum_{i \in \mathcal{N}} c^i |z^i|_0 \quad \Rightarrow \quad \min_{z,r} \sum_{i \in \mathcal{N}} w^i c^i z^i
\]

1. $w^i \leftarrow 1$
2. Iterate:
   1. Solve the linear program;
   2. Update the weights $w^i \leftarrow (z^i + \varepsilon)^{-1}$
3. Terminate after convergence or a maximum number of iterations.
Reweighted $\ell_1$ Minimization

\[
\min_{z,r} \sum_{i \in \mathcal{N}} c^i |z^i|^0 \quad \Rightarrow \quad \min_{z,r} \sum_{i \in \mathcal{N}} w^i c^i z^i
\]

1. $w^i \leftarrow 1$
2. Iterate:
   1. Solve the linear program;
   2. Update the weights $w^i \leftarrow (z^i + \epsilon)^{-1}$
3. Terminate after convergence or a maximum number of iterations.

Energy vs. Traffic

- [Pollakis, et al. `12]
- integer program solver
- basic algorithm
- refined algorithm
Spectrum Allocation (Heavy Traffic)
Spectrum Allocation (Light Traffic)
Post Processing

- Minimizing energy only finds a feasible solution.
- Once the set of active APs is determined, can further minimize average delay as before.
Post Processing (Light Traffic)
Scalability

- Number of variables increases as $O(KN2^N)$
- Infeasible to find optimal allocation for $N \gg 20$.
- To scale to large networks can exploit:
  - Path loss: radio signals cause negligible interference over large enough distances;
  - Small node degrees: typically bounded by a constant
E is the set of network links with non-negligible gain

\[ A_j = \{ i | (i \rightarrow j) \in E \} \]

\[ U_i = \{ j | (i \rightarrow j) \in E \} \]

\[ N_i = \{ \bigcup_{j \in U_i} A_j \} \]

\[ N_2 = A_a \cup A_b \]

\[ N_1 = A_a \]

\[ N_3 = A_b \]

CUHK, April 2016
Local Patterns and Variables

\[ r^j = \sum_{A \subseteq \mathcal{N}} \sum_{i \in A} \mathcal{S}^{i \rightarrow j}_A \mathcal{X}^{i \rightarrow j}_A = \sum_{i \in A_j} \sum_{B \subseteq \mathcal{N}_i} \mathcal{S}^{i \rightarrow j}_B \mathcal{Z}^{i \rightarrow j}_B \]

- Local variables \( z^{i \rightarrow j}_B \) are only defined for links in \( E \) and \( B \) in \( \mathcal{N}_i \).
- Introduce local variables \( y^{i}_B \) defined for \( B \) in \( \mathcal{N}_i \).
- Number of local variables is \( O(N) \).
- Consistency constraint in overlapping neighborhoods:

\[ \sum_{B \subseteq \mathcal{N}_i : B \cap \mathcal{N}_m = C} y^{i}_B = \sum_{B \subseteq \mathcal{N}_m : B \cap \mathcal{N}_i = C} y^{m}_B, \quad \forall i, m \in \mathcal{N}, \forall C \neq \emptyset \]
Relaxed Optimization

- Add previous constraint, optimize over $z, y$
- Relax total bandwidth constraint:
  \[ \sum_{B \subseteq N} y_B \leq 1 \]
- Scale back bandwidth assignments to meet constraint
- Need to satisfy additional alignment constraints
  \( \rightarrow \) strong vertex coloring problem on hypergraph
Delay Example

![Graph showing average packet sojourn time (seconds/packet) vs. average packet arrival rate per user type (packets/second) for different algorithms. The graph includes lines for full-spectrum-reuse + maxRSRP, optimal orthogonal + maxRSRP, and Algorithm 4. The parameters N=25 and K=126 are indicated.]
Concluding Remarks

- Slow resource allocation can exploit spatial traffic variations.
- Centralized optimization
  - Requires gathering traffic statistics across cells
  - Re-optimize periodically
- Network size limited by computational complexity
  - Number of variables increases exponentially
  - Scalability facilitated by optimizing over local neighborhoods