TRAFFIC-DRIVEN RESOURCE ALLOCATION OVER SLOW TIME-SCALES

Michael Honig
Department of EECS
Northwestern University

Joint work with Binnan Zhuang, Dongning Guo.
Heterogeneous Network

How to do interference management?

User Deployed WiFi Access Points/Femtocells/Relays

Operator Deployed Pico cells/Relays

Remote Radio heads

TIW16, Livorno, Italy September 2016
Timescales

Slow time-scale

- 1 millisecond
- 1 second
- 1 minute
- 1 day

opportunistic scheduling
ICIC
eICIC
self-organizing network (SON)
AP placement & configuration
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.

TIW16, Livorno, Italy   September 2016
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

BTS 2

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

BW assigned to BTS 1

BW assigned to both BTS 1 and 2

BW assigned to BTS 2

Total available bandwidth (BW)
Orthogonal Allocation

Traffic for users in cell 1

$\lambda_1 \rightarrow \begin{array}{c}
\text{BW assigned to BTS 1}
\end{array}$

Traffic for users in cell 2

$\lambda_2 \leftarrow \begin{array}{c}
\text{BW assigned to BTS 2}
\end{array}$

$\begin{array}{c}
\text{BTS 1}
\end{array}$

$\begin{array}{c}
\text{BTS 2}
\end{array}$

$\begin{array}{c}
\text{Total available bandwidth (BW)}
\end{array}$

TIW16, Livorno, Italy    September 2016
Full Frequency Reuse

Traffic for users in cell 1

\[ \lambda_1 \rightarrow \text{Traffic for users in cell 2} \]

\[ \lambda_2 \]

BTS 1

All BW assigned to both BTS 1 and 2

Total available bandwidth (BW)

September 2016

TIW16, Livorno, Italy
Partial Sharing

Traffic for users in cell 1

\[ \lambda_1 \rightarrow \]

Traffic for users in cell 2

\[ \lambda_2 \leftarrow \]

Total available bandwidth (BW)

BTS 1

\[ x\{1\} \]

BTS 2

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

\[ x\{2\} \]

BW assigned to BTS 1

Shared BW

BW assigned to BTS 2

TIW16, Livorno, Italy  September 2016
Orthogonal Allocation

Traffic for users in cell 1

\[ \lambda_1 \]

Traffic for users in cell 2

\[ \lambda_2 \]

"I would build a GREAT wall!"

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

BW assigned to BTS 1

BW assigned to BTS 2

Total available bandwidth (BW)
Full Frequency Reuse

Traffic for users in cell 1

\( \lambda_1 \rightarrow \) Traffic for users in cell 2

\( \lambda_2 \)

“Tear down this wall!”

All BW assigned to both
different cells

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

Total available bandwidth (BW)

TIW16, Livorno, Italy  September 2016
Partial Sharing

Traffic for users in cell 1

\[ \lambda_1 \rightarrow \]

Traffic for users in cell 2

\[ \lambda_2 \leftarrow \]

"one country, two systems"

Total available bandwidth (BW)

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

BW assigned to BTS 1

Shared BW

BW assigned to BTS 2

TIW16, Livorno, Italy  September 2016
Partial Sharing

\[ \lambda_1 \rightarrow x \{1\} \]

Traffic for users in cell 1

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

Shared BW

\[ x \{2\} \]

Traffic for users in cell 2

“one country, two systems”

Partition should depend on traffic!

September 2016

TIW16, Livorno, Italy
$N$ Base Stations

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

$\lambda_1 \rightarrow \text{BTS 1}$

$\lambda_2 \leftarrow \text{AP 2}$

$\text{AP 3}$

$\text{AP 4}$

$\text{AP 5}$

$\text{AP 7}$

$\text{BTS 6}$

$\text{x}\{\emptyset\}$

$\text{x}\{1\}$

$\text{x}\{1,2\}$

$\ldots$

$\text{x}\{1,2,3,4,5,6,7\}$

Frequency

TIW16, Livorno, Italy  September 2016
Bandwidth Optimization

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

Rate per Hz:

- BTS 1 transmits
  - $S_1,\{1\}$
  - BW assigned to BTS 1

- BTS 2 transmits
  - $S_1,\{1,2\}$
  - BW assigned to both BTS 1 and 2
  - $S_2,\{1,2\}$
  - BW assigned to BTS 2
  - $S_2,\{2\}$

TIW16, Livorno, Italy  September 2016
Spectral Efficiency

\[ s_{A \rightarrow j}^i = \mathbf{1}_{i \in A} \frac{W}{L} \log \left( 1 + \frac{p_{i \rightarrow j}}{I_{A \rightarrow j} + \sigma^2} \right) \]

- Average powers, channels
- Known to the optimizer
Rate per BTS


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:

\[ r_1 = S_{1,\{1\}} x_{\{1\}} + S_{1,\{1,2\}} x_{\{1,2\}} \]

\[ r_2 = S_{2,\{2\}} x_{\{2\}} + S_{2,\{1,2\}} x_{\{1,2\}} \]
Rate per BTS

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

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

September 2016
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): $t_i = \frac{1}{r_i - \lambda_i}$
Optimization: Backlogged Traffic

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

Subject to:

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

- Convex, \( 2^N - 1 \) variables
- The solution achieves the maximum throughput region.
**Theorem:** The optimal allocation divides the spectrum into at most $N$ segments (instead of $2^N$).

Follows from Carathéodory’s theorem.

**7-BTS example:**

---

TIW16, Livorno, Italy  September 2016
Load to BTS Assignment

BTS 1

BTS 2

Traffic designated for users in region 1

Optimization variables:

$\lambda_1$ $\lambda_2$ $\lambda_3$

Problem: Jointly allocate traffic and bandwidth across base stations.

September 2016

TIW16, Livorno, Italy
Load to BTS Assignment: Notation

- Set of BTSs: $\mathcal{N} = \{1, 2, \cdots, N\}$
- Set of UE groups: $\mathcal{K} = \{1, 2, \cdots, K\}$
- $\lambda_k$: packet arrival rate for group $k$
Load to BTS Assignment: Notation

- $S_{i \rightarrow j}^A$: spectral efficiency of BTS $i$ serving group $j$ under reuse pattern $A$.
- $x_{i \rightarrow j}^A$: spectrum resource used by BTS $i$ to serve group $j$ under reuse pattern $A$.
- $y_A$: fraction of spectrum resources allocated to reuse pattern $A$.
Sub-partition Constraint

\[ \sum_{j \in \mathcal{U}} x_{A}^{i \rightarrow j} \leq y_{A}, \forall A \subset \mathcal{N}, i \in \mathcal{N} \]

\[ \sum_{A \subset \mathcal{N}} y_{A} = 1 \]
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:

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

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

Subject to:

\[ r_i > \lambda_i \]

\[ r_i = \sum_{B \subset \mathcal{N}} s_{i,B} x_B \quad \forall i \in \mathcal{N} \]

\[ x_B \geq 0 \quad \forall B \subset \mathcal{N} \]

\[ \sum_{B \subset \mathcal{N}} x_B = 1 \]
Optimization (Modified)

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

Subject to:

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

\[
x_{B} \geq 0 \quad \forall B \subseteq \mathcal{N}
\]

\[
\sum_{B \subseteq \mathcal{N}} x_{B} = 1
\]
Optimization (Modified)

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

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

Convex for concave \( U \), \( O(KN2^N) \) variables
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
Delay (2 macros, 8 small cells)

- Optimized spectrum allocation, maxRSRP
- Full spectrum reuse, optimized assignment
- Full spectrum reuse, maxRSRP
- Joint optimization

Graph showing average packet arrival rate per user type (packets/second) against average packet sojourn time (seconds/packet). The graph compares different allocation methods.
Scalability

- Number of variables increases as $O(KN2^N)$
- Infeasible to find optimal allocation for $N >> 20$.
- 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

TIW16, Livorno, Italy  September 2016
Local Interference

- Radio signals vanish beyond a certain radius
- Node degrees bounded by a constant
Local Interference

- Radio signals vanish beyond a certain radius
- Node degrees bounded by a constant
- $E$ is the set of links with non-negligible gains
\( \mathcal{N}_i \): APs that potentially interfere with \( i \)
Local Variables

\[ r^j = \sum_{A \subseteq \mathcal{N}} \sum_{i \in A} s^{i \rightarrow j}_A x^{i \rightarrow j}_A = \sum_{i \in A_j} \sum_{B \subseteq \mathcal{N}_i} s^{i \rightarrow j}_B z^{i \rightarrow j}_B \]

- Local variables \( z^i_B \) are only defined for \( i \rightarrow j \) 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 \( O(N) \).
- Need to reduce the global variables \( y_B \) in order to scale.
Relaxed Optimization

- Introduce local variables $y_B^i$, defined for $B$ in $N_i$ only.

- Replace the global constraint

\[
\sum_{B \subseteq N} y_B \leq 1 \quad \text{with} \quad \sum_{B \subseteq N_i} y_B^i \leq 1, \quad i \in N
\]

- Need additional consistency constraint in overlapping neighborhoods:

\[
\sum_{B \subseteq N_i : B \cap N_m = C} y_B^i = \sum_{B \subseteq N_m : B \cap N_i = C} y_B^m, \quad \forall i, m \in N, \forall C \neq \emptyset
\]
Global Allocation: Hyper-graph coloring

- The relaxed solution may violate the global spectrum constraint.
- We seek a global (discrete) carrier assignment.
- The minimum global assignment is equivalent to strong vertex coloring on a hyper-graph.
- We use a heuristic coloring algorithm to yield a global subcarrier assignment.
Approximate vs optimal solution

$n = 12$ APs, $k = 33$ UE groups, $100 \text{ m} \times 100 \text{ m}$

scalable solution using relaxation + coloring

optimal
Delay Example

- 100 APs
- 314 user groups

Color: particular realization
Dashed: proposed method
Solid: maxRSRP, full-spectrum reuse
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