Issues in multiuser diversity

1. Feedback and real-time traffic
2. Traffic prediction and uncertainty
3. Multiple antennas

Mathias Johansson,
Uppsala University
Sweden
The benefit of channel variations

Consider a cellular downlink:
- Multiple users sharing the same resources
- Varying channel quality – ride the peaks!
  - More variations \(\Rightarrow\) more throughput

Multiuser diversity
Limited channel feedback

Consider a downlink using adaptive modulation.

Each timeslot, each mobile feeds back $\log_2(M+1)$ bits indicating which rate the channel supports given a desired BER.

$$r_u \approx \log_2 \left(1 + \frac{\text{SNR}_u}{f(\text{BER}_u)}\right)$$

Example: $M=1$

<table>
<thead>
<tr>
<th>Rate [bits/symbol]</th>
<th>Send 1 if $r_u \geq q_1$</th>
<th>use $q_1$ bits/symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$q_1 = 3$

$r_u < q_1$ use 0 bits/symbol

$q_0 = 0$
Quantization for maximum expected throughput

• The expected throughput with rate thresholds $q_0...q_M$ becomes

$$\langle x \rangle = \sum_{m=1}^{M} q_m \left( \prod_{u=1}^{U} \int_{0}^{q_{m+1}} p(r_u)dr_u - \prod_{u=1}^{U} \int_{0}^{q_m} p(r_u)dr_u \right)$$

  - $U$ is the number of users
  - $p(r_u)$ is the probability that user $u$ can receive with rate $r_u$ at a desired BER.

• 1-bit feedback:

$$\langle x \rangle = q \left( 1 - \prod_{u=1}^{U} \int_{0}^{q} p(r_u)dr_u \right)$$
The downside

As the number of users increases, the throughput becomes extremely sensitive to the choice of \( q \).

- Change from \( q=4 \) to \( q=5 \) for 25 users
  - 3.5 bits/symbol
  - 0.6 bits/symbol
Implications of 1-bit channel feedback

• Theoretically, strict multiuser diversity is not badly affected by limited feedback

• but in practice, an extreme sensitivity to correct quantizations leads to drastic performance drops

• Note also that unfairness increases when feedback is reduced
Possible remedies

• Individual thresholds
  – decrease the sensitivity
  – but optimal individual thresholds depend on other users’ thresholds

• Avoid using strict multiuser diversity

• Increase the channel feedback

We will combine the two first suggestions and at the same time attain short-term fairness
Diversity-Enhanced Equal Access

- A modified fair multiuser-diversity strategy
  - Round-robin tournament:
    - In each time slot transmit to the user with best channel of the users that have not yet accessed the channel.
    - When all users have obtained access, repeat the tournament.
  - Determine individual 1-bit quantizations locally at the mobile terminal
- Ideal for real-time traffic
- Unsensitive to channel quantization errors
Performance

- Note that the scheduler uses a strict multiuser-diversity strategy with a decreasing number of users.
- On average, $U/2$ users compete.
  - Performance will on average equal that of strict multiuser diversity with $U/2$ users.
- Ex: Rayleigh fading, 16 users spread out.
  - Throughput becomes 268% of that of a round-robin scheme with an optimal common 1-bit quantization.
Performance

Multiuser-diversity gain vs. number of users

Normalized system throughput

Users

diff \{ DEEA \} for U=12
2. Traffic prediction and adaptation

- In theory, better performance can be obtained by scheduling over several time slots
  - Particularly with QoS constraints
- Requires channel prediction over longer periods
  - Should average criterion over pdf for channel to account for higher uncertainty
- Requires traffic prediction
  - "Always data to send" unrealistic assumption
Scheduling under uncertainty

• Minimize the expected total buffer contents after the scheduled horizon
  – Gives maximum expected throughput

• Constrain assignments as in DEEA to satisfy delay constraints

\[
\langle L \rangle = \sum_{u=1}^{U} \sum_{n_u=0}^{\infty} \sum_{x_{ut}=0}^{\infty} p(n_u|I)p(x_{ut}|I)g\left(S_u + n_u - \sum_{t=1}^{T} x_{ut}\right)
\]

\[g(x) = x \text{ if } x > 0, \text{ otherwise } g(x) = 0\]
Traffic prediction

- We wish to determine \( p(n|I) \) given \( I = \) past arrival statistics
- Imagine using histograms
  - Too few observations in comparison to possible inflow sizes
- Instead, partition the inflow-axis into a number of ’bins’
  - Count arrivals within each bin
  - Adapt the bin size to obtain high resolution at intervals of high intensity and lower elsewhere
Traffic prediction – Bin probabilities

Using K bins and letting

– \( m_k \) = past number of arrivals of size within bin \( k \)

– \( M \) = total number of observations

we have (after some calculations...)

\[
p(n \in \text{bin } k \mid m_k, M, I) = \frac{m_k + 1}{M + K}
\]
Traffic prediction – Adaptation

• Based on the bin probabilities, how do we adapt the bin positions and sizes?
  – Wish to have a quantized distribution which is as close to the exact distribution as possible.

• Formally, we wish to maximize the mutual information between the two distributions

Theorem:
Maximizing the mutual information is equivalent to maximizing the entropy of the bin probability distribution.
Traffic prediction – Adaptation

Proof:

\[ I(k, n) = \sum_{k=1}^{K} \sum_{n=n_{min}}^{n_{max}} p(nk) \log \frac{p(nk)}{p(n)p(k)} \]

\[ = \sum_{k=1}^{K} \sum_{n=n_{min}}^{n_{max}} p(nk) \log \frac{p(k \mid n)}{p(k)} \]

\[ = - \sum_{k=1}^{K} \sum_{n \in \text{bin } k} p(n \mid k)p(k) \log p(k) \]

\[ = - \sum_{k=1}^{K} p(k) \log p(k) \]
Traffic prediction

• The optimum bin partition is adapted according to the M most recent arrivals:
  – Assume a uniform probability distribution within each bin,
    \[ p(n) = \frac{\text{bin probability}}{\text{bin width}} \]
  – Redistribute the bins so that each bin has equal probability mass (=max entropy)
  – Approximate low-complexity solution requires single sweep over the possible arrival sizes.
Traffic prediction – Simulation set-up

- K = 6 bins,
- M = 100 observations between bin updates
- Min arrival rate=0, Max=100 [bit/time unit]
- Arrivals generated as
  - 50% 1-bit packets,
  - 50% 7-bit packets
  (switching between 2 fixed rates with equal frequency)
Traffic prediction – Results

After data block 1

After data block 2

After data block 3

After data block 4

After data block 5

i

0 10 20 30 40 50 60 70 80 90 100
3. MIMO-Multiuser Diversity

• With a single channel, capacity-optimal schedule seds to one user at a time (Knopp&Humblet -95)
  \[ C \sim \log(\log U) \]

• Recently, the Shannon capacity for the MIMO case has been found (Caire&Shamai 2001, Vishwanath&Jindal&Goldsmith 2002, Viswanath&Tse 2002)

• The capacity-optimal scheme requires full channel knowledge at the transmitter and is achieved by Costa-precoding (extremely complex)

• Single-user transmission no longer optimal!
MIMO-Multiuser Diversity

- Consider M Tx antennas, and 1 Rx antenna and full channel knowledge at transmitter and receiver
- Then, capacity for Gaussian MIMO broadcast channel for large U is
  \[ C \sim M \log(\log U) \]
  (Sharif & Hassibi -03, submitted)
- Sharif & Hassibi proposes using M random beams where each user feeds back best beam and SINR
  \[ \Rightarrow \text{approaches capacity when } U \to \infty \]
A simple MIMO-multiuser approach

• Transmit to different users on each antenna

• Let each user feedback best SINR and best antenna

• On each antenna, transmit to the user with highest SINR
A simple MIMO-multiuser approach

- Assuming $M = 2$ Tx antennas, and that noise is small compared to the interference from the other antenna, the rate is

$$R_u \leq \log \left( 1 + \frac{|h_{1u}|^2}{\sigma^2 + |h_{2u}|^2} \right)$$

$$\approx \log \left( 1 + \frac{|h_{1u}|^2}{|h_{2u}|^2} \right)$$
A simple MIMO-multiuser approach

• Assume independent, identically-distributed flat Rayleigh fading on both antennas. (Exponentially distributed \(|h|^2\)).

• The distribution of a ratio of two exponentially distributed numbers with the same mean is independent of that mean.

Thus, the performance will not depend on the average channel gain!

– Unfairness no longer a big problem
Performance

• The expected throughput depends on the probability for having high gain on one antenna, low gain on the other.

\[ R_u \approx \log \left( 1 + \frac{|h_{1u}|^2}{|h_{2u}|^2} \right) \]

• More users ⇔ Higher probability!

• Compare the expected throughput of this scheme to that of
  – always sending only on antenna 1,
  – sending only to the best user on the best antenna (equal to one-antenna case with twice as many users)
Performance

Expected throughput for 30 users (independent Rayleigh)

30 users

- MIMO Interference
- Best Tx
- Tx 1 always
Performance

Expected throughput for 16 users (independent Rayleigh)

- 16 users
- 21 dB

Expected throughput for 4 users (independent Rayleigh)

- 4 users
- 11 dB
Multiuser-diversity gain

Expected throughput vs number of users

- MIMO Interference
- Best Tx; E(SNR) = 15 dB
- Tx 1 always; E(SNR) = 15 dB
- Tx 1 always; E(SNR) = 20 dB
- Best Tx; E(SNR) = 20 dB

20 dB
15 dB
Concluding remarks

- Real-time traffic and multiuser diversity is compatible for large $U$
  - Reason is $\log(\log U)$ behavior of rate
- Traffic prediction will be required if scheduling over longer time horizons
  - presented new adaptive method
- Multiuser-MIMO requires transmitting more than one data stream at a time to approach capacity
  - simple interference-channel approach seems promising