Fast Handovers for
Heterogeneous Cellular Networks

Shivendra S. Panwar
Joint work with Sundeep Rangan, Elza Erkip, Pei Liu,
Ayaskant Rath and Sha Hua
Outline

- Challenges in Today’s Cellular Systems
  - Future Directions
  - Heterogeneous Networks
- Fast Handover in Networks with Femtocells
  - Legacy Handover Issues
  - Fast Handover
  - Benefit Analysis
- The Urge to Merge
  - Ideas
  - Preliminary Analysis
Outline

- Challenges in Today’s Cellular Systems
  - Future Directions
  - Heterogeneous Networks
- Fast Handover in Networks with Femtocells
  - Legacy Handover Issues
  - Fast Handover
  - Benefit Analysis
- Urge to Merge
  - Ideas
  - Preliminary Analysis
Global Mobile Data Traffic Forecast

- Cisco predicts mobile data traffic to grow 26-fold by 2015 from 2010.

Source: Cisco VNI Mobile, 2011
Future of Data Traffic

- The biggest driver for the traffic increase will come from video traffic.
  - Video could account for roughly 64% of all mobile data traffic in 2013. It is currently 39%.

- Handsets and laptops with speeds higher than current 3G speeds will account for 80% of all mobile traffic by 2013.

- Emerging applications including video conferencing, telemedicine, interactive gaming and mobile education systems.
Traffic and Revenue have decoupled

- Network operators cannot afford to support such explosive demand.
  - Traditionally, operators have been adding BSs and spectrum to relieve congestion.
  - Would work well if revenue scales linearly with demand.
  - This is not the case for data.
    - High usage of network resources.
    - Revenue per bit: SMS >> Voice >> Data.

Addressing the Data Crisis

- **Spectrum Efficiency**
  - Femtocells / WiFi traffic offloading.
  - Interference Mitigation.
  - Cooperative Relaying.
  - Distributed Antenna Systems
  - Coordinated Multipoint (CoMP)

- **Video optimization for wireless**

- **More Spectrum**
  - Millimeter wave and sub-THz wireless communications.
Evolution towards Coherent Heterogeneous Integrated Cellular (CHIC) Networks (joint work with Sundeep Rangan)

- Cellular Networks today, are becoming heterogeneous in terms of applications, devices and infrastructure.

- The intention is to find solutions that enable massively heterogeneous cellular networks to
  - realize their potential by unlocking hidden capacity, alleviating congestion and enabling innovation.
  - mitigate risks of poor mobility support, inefficient interference management and unstable nested control.
Emerging Cellular Model

- Open Access
- Heterogeneous
- Scalable
- Self-Organizing
- De-centralized
## Heterogeneous Network Vision

<table>
<thead>
<tr>
<th>Network Model</th>
<th>Density/Capacity</th>
<th>Mobility Support</th>
<th>Openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current, vertically integrated macrocell networks</td>
<td><strong>Poor</strong></td>
<td><strong>Excellent</strong> Full high-speed handovers.</td>
<td><strong>Poor</strong> Closed, centralized architecture.</td>
</tr>
<tr>
<td></td>
<td>Cannot scale in a cost effective manner.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macrocells with femtocells and WiFi hotspots</td>
<td><strong>Under Study</strong></td>
<td><strong>Poor</strong> Typically, nomadic only.</td>
<td><strong>Good</strong> Capacity offload with non cellular connectivity.</td>
</tr>
<tr>
<td></td>
<td>Interference issues under active study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coherent Heterogeneous Integrated Cellular Networks</td>
<td><strong>Good</strong></td>
<td><strong>Good</strong> High-speed mobility.</td>
<td><strong>Good</strong> Open, non-cellular airlink with backhaul access.</td>
</tr>
<tr>
<td></td>
<td>Resolve interference issues for dense, high capacity coverage.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
WIRELESS INTERNET CENTER FOR ADVANCED TECHNOLOGY
NSF INDUSTRY/UNIVERSITY COOPERATIVE RESEARCH CENTER

Heterogeneity and Mobility

From Gateway Architecture to Intelligent Edge

(a) Current “evolved packet core” (EPC)  
(b) Proposed “edge” network
Benefits of Heterogeneity

- Evolved Packet Core (EPC) is motivated by several considerations:
  - It offers a consistent IP point of attachment that makes user mobility transparent to the public internet.
  - Handling all mobility procedures within the EPC enables highly optimized procedures for handover, access control and paging with significant cross-layer design.
  - It simplifies the design of terminal and centralized processing with terminals only providing measurement reports.

- Why switch to Intelligent Edge?
  - Scalability
  - EPC handover is designed generally assuming the air interface to be the bottleneck and hence relies on extensive, fast messaging in the core network.
  - EPC may have poor performance in heterogeneous networks where access points could have limited backhaul and/or significant delay to the core network.
Benefits of Heterogeneity (contd.)

- Cellular standards have already begun considering support for providing an IP point of attachment directly to the BSs.
  - Selective IP Traffic Offload (SIPTO)
  - Local IP Access (LIPA)
- If mobility these networks can support mobility, edge networking can provide the basis for a more scalable architecture without an expensive dedicated core network and enable integration of more varied backhaul access technologies.
  - Services including caching and content delivery networks (CDNs).
  - Adaptive video transcoding and low-latency network virtualization and cloud computing.
Robust Application Layer Mobility

Make-before-break handovers that support simultaneous physical layer connections to both source and target BSs.

1. SIP-based mobility
2. Transport protocols such as Stream Control Transmission Protocol (SCTP) and Multipath TCP.
Local Mobility via Relaying

- Switching IP point of attachment can be highly time consuming during fast handovers.
- Over-the-Air, PHY-layer relaying may be more appropriate.
- Multiple streams of data can be forwarded to the relay eNodeB to assist transmissions or re-transmissions.
Cooperation in Heterogeneous Wireless Networks
Joint work with Elza Erkip and Pei Liu

- Peak data rate has increased dramatically over the years, but problems still exist at the cell edge.
  - Stations at the edge transmit at a much lower data rate than stations at the center (10:1 for WiFi).
  - Packet transmission for a station at the edge takes longer.
  - Higher interference level at the cell edge.

- Cooperative communications allow a third node (relay) to help.
  - Relays process this overheard information and forward to destination.
  - Network performance is improved because edge nodes transmit at higher rate thus improving spectral efficiency.
  - Candidate relays: Mobile user, pico/femto cell BS, fixed relays, etc.
  - Incentives: Throughput, power, interference.
Relays in Commercial Systems

- Cooperative/multihop communications have already been adopted in next generation wireless systems.
  - **IEEE 802.11s**
    Enables multihop and relays at MAC layer, does not provide for joint PHY-layer combining.
  
  - **IEEE 802.16j**
    Expands previous single-hop 802.16 standards to include multihop capability. Integrated in IEEE 802.16m draft.
  
  - **3GPP LTE**
    Cooperative multipoint is supported with joint transmissions and receptions for cost-effective throughput enhancement and coverage extension.
Cooperation and Heterogeneity: A Good Fit

- Cooperation performs much better if the number of relays is large.
  - In macrocell based deployment, the number of operator deployed relay stations is limited.
  - In traditional networks, performance gain for cooperation is limited unless user cooperation is enabled.
  - User cooperation gives rise to problems like high battery consumption, synchronization, security and lack of incentives.

- Proliferation of pico/femto BSs will provide a large number of “good enough” relays.
  - No battery consumption problem.
  - Easier to synchronize: stationary, backhaul connection, better radio design.
  - More secure because they are part of operator’s network.
Cooperative MIMO for Het Nets

- For highly mobile users and for users without femtocell coverage, cooperative MIMO
  - enables fully *opportunistic* use of all available surrounding radios.
  - increases network capacity and helps reduce coverage holes.
Comparison of Relaying Schemes

Limitations of previous cooperative methods:

- **Single Relay**: low spatial diversity gain.
- **Multiple Relays**: consume more bandwidth resources when several relays forward signal sequentially.

Alternatives:

- **Distributed Space Time Coding (DSTC)**.
  - Recruit multiple relays to form a virtual MIMO.
  - Each relay emulates an indexed antenna.
  - Each relay transmits encoded signal corresponding to its antenna index.

- **Pro**: Spatial diversity gains

- **Cons**:
  - Tight synchronization required.
  - Relays needed to be indexed, leading to considerable signaling cost.
  - Global channel state information needed.
  - Good DSTC might not exist for an arbitrary number of relays.
  - Unselected relays cannot forward, sacrificing diversity gain.
Robust Cooperative MIMO

- Randomized cooperation strategies provide powerful PHY-layer coding techniques that
  - alleviate the previous problems and allow robust and realistic cooperative transmission with multiple relays.
  - randomize distributed space time coding (R-DSTC) for diversity.
  - randomize distributed spatial multiplexing (R-DSM) for spatial multiplexing.

- Highlights of randomized cooperation:
  - Relays are not chosen a priori to mimic particular antennas.
  - Multiple relays can be recruited on-the-fly.
  - Relays are used opportunistically according to instantaneous fading levels.
  - Signaling overheads and channel feedback are greatly reduced.
  - Performance comparable to centralized MIMO is attained.
R-DSTC Performance (WiFi)

- Underlying orthogonal STBC codeword size: 2, 3, 4.
- PHY-layer rates: 6, 9, 12, 18, 24, 36, 48, 54.
- BPSK, QPSK, 16-QAM, 64-QAM, Convolutional Code $\frac{1}{2}$, $\frac{2}{3}$, $\frac{3}{4}$.
- 20MHz bandwidth.
- Contention window: 15-1023.
- Transmission power: 100mW.

CoopMAX: A Cooperative Relaying Protocol for WiMAX

- CoopMAX enables robust cooperation in a mobile environment with low signaling overheads.
- It is robust to mobility and imperfect knowledge of channel state.
- Simulation shows 1.8x throughput gain for a single cell with mobility and 2x throughput gain for multicell deployment.

Performance

- Our results demonstrate that R-DSM scheme delivers MIMO system performance.
  - Average data rate for the second hop (relays-destination link) scales with the number of relays.
  - For direct transmissions, the peak data rate is supported at a short range.
  - R-DSM can increase the number of stations that can transmit near the peak data rate.

Business case for Femto-relays

- There are 1000 customers per macrocell
- Assume 10% penetration of “free” femto-relays
- This implies 100 femto-relays per macrocell
- This doubles cell capacity using the relay function
- Assume each femto-relay costs $100
- Cost per customer = \(100 \times 100 / 1000 = $10\)
- $0.50 charge per month will recover this < 2 years
- Total US cost (300 million customers) = $3 billion
- AT&T’s purchase price for T Mobile for roughly similar capacity gain: $39 billion
Outline

- Challenges in Today’s Cellular Systems
  - History
  - Future Directions
  - Heterogeneous Networks

- Fast Handover in Networks with Femtocells
  - Legacy Handover Issues
  - Fast Handover
  - Benefit Analysis

- The Urge to Merge
  - Ideas
  - Preliminary Analysis
Architecture with Femtocells

- Macrocell Base Stations [(e)NB] are part of the Mobile Core Network [MCN].

- Femtocell Base Stations [H(e)NB] connect to the MCN via the public internet.

- There is an additional delay for all control (and data) messages to reach the MCN from the H(e)NB.

- The additional delay is of the order of 100s of milliseconds.

- Thus handovers involving femtocells are slower.
Handover in Cellular Networks

- UE measures signal strengths on its associated channel, and on broadcast channels periodically.
- UE associates with the base station (Macro or Femto) with the strongest signal.
- Handover = Change of association
  - Occurs when the signal from a new base station is stronger than that from the associated base station.
  - Must be as seamless as possible.
- Types of handover:
  - Handout: Femto to Macro
  - Handin: Macro to Femto
  - Handover: Femto to Femto
  - Macro to Macro handover is already highly optimized.
- Legacy Handover procedures were originally designed with only Macrocells under consideration.
Legacy Handover Procedure applied to Femtocells

- **Handover Decision**
  - Handover Request
  - Handover Response
  - Admission Control
  - Handover Request
  - Handover Response

- **Packet Data**
  - Buffer Packets from SGW

- **Detach Src Attach Tgt**

- **Handover Confirm**
  - Path Switch Request
  - Path Switch Response

- **Switch Path**
  - UP Update Req
  - UP Update Rsp

- **Release Resource**
  - Packet Data
Legacy Handover – Points to Note

- UE receives weak signal at least from the time Handover Decision is made until it detaches with the source BS.

- UE cannot receive downlink data from the time it receives Handover command until forwarded data from source BS is received by target BS.

- As many as four control plane message exchange cycles (shown in blue) involve message traversal over the internet (shown in bold arrows).

- During the handover process, data from SGW reaches the target BS via the source BS (shown in red), which is two sequential traversals over the public internet.

- Admission control process, and Path Switch process are tightly coupled with the Handover process.
Fast Handover

- Problems with legacy Handover in networks with dense deployment of open Femtocells
  - More Frequent.
  - Slower.
  - Sometimes fruitless.

- Signal Strength Measurement Reports from the UE can be used to:
  - Choose association.
  - Estimate Speed of UE.
  - Estimate Location of UE.
Fast Handover (contd.)

- **UE Status:**
  - *Low speed State:* When the estimated speed of the UE is below a threshold $\Omega$.
  - *Flying State:* When the estimated speed of the UE is above $\Omega$.
    - Only handouts allowed.

- **UE in Handover Proximity Range:** When the signal strength difference between the strongest associable signal and another BS is below a threshold $\Delta$.
  - $\Delta$ can be network-wide or UE specific
Fast Handover – Cellular Regions

- **Strict Association Region**: Difference between the strength of the best and the next best signals from BSs is higher than $\Delta$.

- **Handover Proximity Region**: Signals from one or more BSs closer than $\Delta$ to the best signal.
  - The BSs whose signal strength is within $\Delta$ of the best signal are said to be *in proximity* of the UE.
Fast Handover Features

- When the UE is in Strict Association Region: No change.
- When the UE is in Handover Proximity Region:
  - Associated with one BS (same as legacy handover).
  - *Potential* near future association with more than one BS.
  - Application layer data buffered at all proximate BSs.

- Elements of the Handover procedure decoupled into:
  - Handover Proximity Add: When a new BS enters the proximity of the UE.
  - Handover Proximity Release: When a BS in the UE’s proximity leaves it.
  - Handover.

- Transparent to the UE (no change from the legacy Handover).
- Effectively, actual Handover becomes much faster for the UE.

- Soft modification in the MCN design required.
- Higher resource consumption.
Handover Proximity Add/Release – Points to Note

- These procedures are triggered well before/after the actual handover process.

- When a BS is added to the proximity of the UE, data packet flow from SGW is duplicated to the new BS, which buffers the packets.
  - This can be done using multi-flow session provision in SCTP.

- When a BS is released from the proximity of the UE, data packets buffered are discarded.
  - This is the cost (wasted resource) we pay for higher speed handover.
Fast Handover Procedure

1. **Measurement Reports** from **UE** to **Src H(e)NB**
2. **Handover Decision** based on received reports
3. **Handover Request** from **Src H(e)NB** to **H(e)NB-GW**
4. **Handover Response** from **H(e)NB-GW** to **Src H(e)NB**
5. **Packet Data** transferred from **Src H(e)NB** to **Tgt H(e)NB**
6. **Handover Command** from **Tgt H(e)NB** to **UE**
7. **Detach Src Attach Tgt** from **UE**
8. **Buffer Packets from SGW**
9. **Status Transfer** from **SGW** to **H(e)NB-GW**
10. **Handover Confirm** from **Tgt H(e)NB**
11. **Switch Marker**
12. **Discard “Stale” Buffered Packets**
13. **Packet Data** transferred from **Tgt H(e)NB** to **H(e)NB-GW**
14. **Proximity Check**
15. **Handover Request** from **MME**
16. **Handover Response** from **Tgt H(e)NB**
17. **Packet Data** transferred from **Tgt H(e)NB** to **SGW**
Fast Handover Procedure – Points to Note

- Handover Decision happens exactly when it happens in legacy handover procedures.

- Admission control process is replaced by much simpler and quicker Proximity Check process when ensures the target BS is already in the proximity of the UE.

- UE receives weak signal from the time Handover Decision is made until it detaches from the source BS.
  - This duration is shorter than in legacy Handover procedure.

- UE cannot receive downlink data from the time it receives Handover command only until a single small Switch Marker packet is received by the target BS from the source BS.
  - This duration is shorter than in legacy Handover procedure.
  - Target BS already has buffered data packets and is ready to continue session with UE.
  - No data packet forwarding from source to target BS.

- No Path Switch process.
Simulation Model

- One macrocell: Range: 1km
- 100 Femtocells: Range: 50m (approx.)
- Proximity Threshold: 15m (approx.)
- Random Walk Mobility Model
- Maximum Speed: 35mph

- Measurement Report Frequency: 10/s
- Message traversal over MCN: 10ms
- Message traversal over the internet: 210ms
- Other functions: <10ms

- Total Duration: 1 hour (approx.)
Area Map
Initial Results

- Total number of handovers: 158
- On an average, fast handovers consume 53% less time compared to legacy handovers.
- On an average, maximum speed sustainable by fast handover is 38% higher than that sustainable by legacy handover.
- Time spent in transmitting discarded data: 10.37s per handover
Potential Tweaks

- **Point of data multiplexing can be changed to H(e)NB-GW.**
  - Depends on where H(e)NB-GW is actually placed in the network.
  - Can make the fast handover procedure transparent to all components of the network except H(e)NB, H(e)NB-GW and MME.
  - Save resources.

- **Direct handoff signaling from umbrella (e)NB.**
  - Depends on feasibility of over-the-air link between the macro and femto BSs.
  - Further improve Handover Speed.

- **Handover Proximity Region based on Mobility Prediction.**
  - Predict user mobility based on measurement data.
  - Reduce the number of BSs in proximity of the UE.
  - Save resources.
Outline

- Challenges in Today’s Cellular Systems
  - History
  - Future Directions
  - Heterogeneous Networks
- Fast Handover in Networks with Femtocells
  - Legacy Handover Issues
  - Fast Handover
  - Benefit Analysis
- The Urge to Merge
  - Basic concept
  - Preliminary Analysis
The Urge to Merge

- What if cellular operators could merge and pool resources?
  - Share infrastructure.
  - Share spectrum.
How can BSs be shared

- **Traditional Roaming**
  - Only works when no connection is available to the assigned operator (e.g., connect to AT&T when signal from T-Mobile is non-existent).
  - Stringent constraints and high costs.
  - How about a more general case?

- **Sharing the base stations**
  - If users can freely access the BSs of either operator using the *closest-first* rule.
  - Network capacity can be improved.
Co-located Cell Towers

[Scenario 1]

- Base Stations from two operators are co-located on the same cell tower.
  - Multiplexing gain by opportunistically connecting to the BS with lower traffic load.
Cell Towers retain spectrum; users roam

Scenario 2

- Base Stations from two operators (red and blue using yellow and light blue spectrum respectively) are maximally distributed with each other.

- A frequency reuse-2 system
- Halve cell size
- Reduces inter-cell interference
- Reduced SINR
Back-of-the-Envelope Calculation

- Downlink edge user power gain (point A in original cell, B in new cell)
  - The distance reduces to $\frac{\sqrt{2}}{2}$ of the original. When the path loss exponent is 4, received power gain is 6dB.
  - Edge user boost from BPSK $\frac{1}{2}$ to QPSK $\frac{3}{4}$. Spectrum efficiency increases by 3x. Interference remains at almost the same level.

- 3x theoretical gain for edge users in Scenario 2.

- Downlink user average
  - Average user signal strength improved by 6.02dB.
  - Average user interference reduced by 0.91dB.

<table>
<thead>
<tr>
<th>Modulation</th>
<th>Coding Rate</th>
<th>Receiver SNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPSK</td>
<td>$\frac{1}{2}$</td>
<td>3.0</td>
</tr>
<tr>
<td>QPSK</td>
<td>$\frac{1}{2}$</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>$\frac{3}{4}$</td>
<td>8.5</td>
</tr>
</tbody>
</table>

IEEE 802.16e
More Realistic Cells

- In a hexagonal layout
  - Downlink edge user performance
  - Performance of points A, C and E remains the same. Points B, D and F move to the center of the new cell and have better performance.

- Generally, users in green area (edge users) achieve performance gain.
Further Improvement

- Without adding more hardware, Scenario 2 triples edge network capacity.

- What happens if new hardware were added to the newly acquired cell towers?

Scenario 3: Aggregate spectrum

- Adding more infrastructure
  - Aggregate the two spectrums at each BS location.
  - 2x gain for aggregate capacity.
Preliminary Results

- Real BS locations in Washington DC area (20km x 20km).
- 16 towers from one operator and 13 from another.
- Randomly scattered users (100 per cell on average)

<table>
<thead>
<tr>
<th>User’s Avg. Rate</th>
<th>No Cooperation</th>
<th>Scenario 2 (users can flexibly roam)</th>
<th>Scenario 3 (aggregate spectrum)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical Model</td>
<td>192.62kbps</td>
<td>279.20kbps</td>
<td>386.36kbps</td>
</tr>
<tr>
<td>OFDMA-based Simulation</td>
<td>270.08kbps</td>
<td>402.74kbps</td>
<td>570.98kbps</td>
</tr>
</tbody>
</table>
Summary

- Introduced CHIC: “HetNets plus”
- Femtorelays
- Fast handover for CHIC/HetNets
- Some interesting engineering aspects of the AT&T/T-Mobile merger!
Questions

Thank You