Heterogeneous Cellular Networks

A solution to cellular data crisis?

Shivendra S. Panwar, Professor
Polytechnic Institute of NYU

Tutorial for Wireless Symposium @ Virginia Tech
Syllabus

• A Brief History of the Cellular System

• The Cellular Data Crisis

• Evolution Towards a Coherent Heterogeneous Integrated Cellular (CHIC) Network
  – Network Architecture
  – Interference Management & Macro-Femto Spectrum Sharing
  – Mobility Support

• Cooperative Relaying
  – Cooperative Data Transmission
  – Cooperative Video Transmission
  – Cooperative Communications for the Battlefield

• Coordinated MultiPoint Transmission (CoMP)

• The Urge to Merge: Pooling Spectrum Together

• Adapting Video for Wireless Networks
  – Scalable Video Coding
  – Perceptual Video Quality

• Millimeter Wave and Sub-THz Wireless Communications
A Brief History of the Cellular Systems

• 1G - Analog modulation and mobility support
• 2G - Digital modulation using TDMA. The radio access network (RAN) is designed only for voice
• 3G - CDMA based transmission and the RAN is still designed for voice with limited data support
• 4G – OFDMA based transmission and is designed to support mobile Internet. The first priority is not voice anymore.
Acknowledgments to WICAT colleagues for use of their slides

- Pei Liu
- Sundeep Rangan
- Yao Wang
- Elza Erkip
Cellular Systems Structure

• Since their inception, cellular networks have been designed, built, and operated as vertically integrated systems
  – Operator owns network and spectrum
  – Devices are built with closed architecture
  – Applications are well defined (voice/SMS)

• It allows a very tight coupling between all network elements

• Very efficient in term of mobile voice support
How did the current cellular network scale with demand?

• Increase in network capacity during 1985-2008 was 10,000
  – Cell splitting  \times 400
  – More Spectrum  \times 3.4
  – Efficient Technology  \times 7

Source: David Goodman, personal communication
Cell Splitting

• By deploying more BS’s having more than one cell site cover a particular amount of geography

• This process creates more coverage and capacity in high demand urban areas

• Very high deployment and OPEX cost

• Currently underway in major US cities to keep up with the high traffic volume of smartphones
Technology Evolution

- Multiple access schemes
- Wider transmission bands
- Efficient digital modulation and coding makes 6 bits/sec/Hz possible
- MIMO technology improve the capacity by another factor of 2
Spectral Efficiency is Approaching Limits

Source: LTE Advanced: Heterogeneous Networks, Qualcomm, February 2010
FCC Auction 2008
$19,000,000,000

Revised 700 MHz Band Plan for Commercial Services

LOWER 700 MHz BAND
(TV CHANNELS 52-59)

UPPER 700 MHz BAND
(TV CHANNELS 60-69)
The Data Crisis

- Wireless data traffic volume is increasing exponentially
  - Driving by data hungry smartphones and new applications on them
    - iPhone typically generates 30 times the data traffic of a basic-feature phone
  - Demand for mobile computing systems

* Monthly basic mobile phone data traffic
Source: Cisco VNI Mobile, 2011

Shivendra S. Panwar
2011 Wireless Symposium @ Virginia Tech
Global Mobile Data Traffic Forecast

- Cisco predicates mobile data traffic to grow 26-fold by 2015 (from 2010)
Global Mobile Data Traffic Forecast

• The biggest driver for the traffic increase will come from video traffic
  – Account for roughly 64% of all mobile data traffic in 2013 (39% currently)

• Handsets and laptops with speeds of 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 & Revenue Decoupled

- Network operators cannot afford to support such explosive demand
  - Traditionally operator have established new BS to relieve congestion for voice
  - Would work well if revenue scales linearly with demand
  - However, that’s not the case for data
    - High usage for network resources
    - Per bit revenue: SMS >> Voice >> Data

Source: Unstrung Insider, Mobile Backhaul and Cell Site Aggregation, Feb 2007

Shivendra S. Panwar
2011 Wireless Symposium @ Virginia Tech
Addressing the Data Crisis

• Spectrum Efficiency beyond Shannon Coding & MIMO
  – Femtocell/ WiFi traffic offloading
  – Interference Mitigation
  – Cooperative Relaying
  – Coordinated multipoint Tx/Rx (CoMP)

• Video Optimization for Wireless
  – Scalable video codec
  – Perceptual video quality
  – Context aware compression and adaptation

• More Spectrum
  – Millimeter wave and sub-THz wireless communications
Evolution towards Coherent Heterogeneous Integrated Cellular (CHIC) networks

- The cellular network are becoming *heterogeneous in terms of applications, devices, and infrastructures*

- The network architecture *cannot remain vertically integrated and should allow open access*

- The intention is to find solutions that enable massively heterogeneous cellular networks
  - Realizing the potential of these networks for *unlocking hidden capacity, alleviating congestion, and enabling innovation*
  - Mitigating the risks of these networks in terms of poor mobility support, inefficient interference management, and unstable nested control
Emerging Cellular Model

• Open Access
• Heterogeneous
• Scalable
• Self-organizing
• De-centralized
Femtocells: A Personal Base Station

- Small, personal base station in customer’s premises, but in operator’s spectrum
- Possibility for greatly increased capacity at low cost
  - Replace large macrocell tower and related RF equipment with small access point
  - No operator deployment or maintenance.
  - Offload traffic to subscriber’s backhaul

## 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 network</td>
<td>Poor: Cannot scale in cost effective manner</td>
<td>Excellent: Full high-speed handovers</td>
<td>Poor: Closed, centralized architecture</td>
</tr>
<tr>
<td>Macro with femto &amp; WiFi hotspots</td>
<td>Under study: Interference issues under active study</td>
<td>Poor: Typically nomadic only</td>
<td>Good: Capacity offload with non-cellular connectivity</td>
</tr>
<tr>
<td>Coherent heterogeneous cellular</td>
<td>Good: Resolve interference issues for dense, high capacity coverage</td>
<td>Good: High-speed mobility</td>
<td>Good: Open non-cellular airlink and backhaul access</td>
</tr>
</tbody>
</table>
Interference in Het Nets

- Interference coordination in CHIC networks is particularly challenging
  - Dense, small cell sizes worsen the interference condition, reduce traffic smoothing and aggregation, and heighten the need for fast handovers
  - Unplanned deployments increase the dynamic range of interference variation and power disparities
  - Diverse ownership of control and configuration introduces nested loops of resource control spanning large timescale and spatial ranges
Interference in Het Nets

- Distributed coordination of the air-interface in unplanned deployments with strong and varied interference

Standard macrocellar model with three-way sectorization

Femtocell “hotspots” interfere with one another and with the macro overlay

A CHIC network where mobiles connect freely to a “sea” of femtocells
Stronger and Varied Interference

- Traditionally, MS always connects to the strongest signal
- Due to restricted association, a small but significant fraction of links in the femto-cellular model experience very strong interference
Cellular Interference Control Today

Power Control & Reuse 1

• Since advent of CDMA, power control + reuse 1 is the basis for cellular interference management:
  – All links transmit on entire bandwidth (reuse 1).
  – Each link sees average of interference from many small sources.
  – Rate adaptation, power control used to adapt to interference.

• Many effective algorithms:
  – Problems are generally convex
  – Solutions implementable in current standards
Lessons from the Gaussian Interference Channel

- Reuse 1 and power control has been the dominant model for cellular systems, esp. since CDMA
- But, stronger interference conditions in Het Nets requires more sophisticated methods.
Two Commonly Used Strategies for Interference Coordination

**Power control with Reuse 1:**
- Links transmit at same time & freq
- Tradeoff link conditions via power control
- Basis of 3G and 4G cellular systems
- Works well when interference is low.

**Orthogonalization:**
- Links transmit at different times or freq
- Tradeoff link conditions via bandwidth allocation
- Basis of 802.11 standards and earlier 2G cellular systems.
- Works well when interference is high or unpredictable.
The New Interference Environment
Femtocells & Heterogeneous Networks

• Next-gen networks scaling capacity via heterogenous networks:
  – Mixtures of femtocells & relays

• A new interference environment:
  – Deployment is unplanned.
  – Cell selection may be restricted (backhaul, access restrictions)
  – Limited ability to coordinate interference control

Interference scenarios in 3GPP 25.820, HNB study item

Shivendra S. Panwar
2011 Wireless Symposium @ Virginia Tech
Dynamic Interference Coordination

- Dynamic orthogonalization provides a distributed method for coordinating interference
  - Greatly improved capacity
  - Distributed implementation

Source: S. Rangan, R. Madan, Belief Propagation Methods for Intercell Interference Coordination, July 2010
Belief Propagation Solution

\[ p_i \leftarrow j(x_j) = \prod_{s \neq i} p_s \rightarrow j(x_j) \]

\[ p_i \rightarrow j(x_j) = E\left[ \exp\left( u_f(x_i, z_i) \right) \mid x_j \right] \]

- Iterative message passing algorithm
  - Widely used in coding, non-Gaussian estimation, machine learning

- Pass “beliefs” along edges of graphs representing estimates of the marginal distribution

- Natural distributed implementation for wireless.
Belief Propagation Simplifications

\[ p_i \stackrel{j(x_j)}{\leftarrow} \prod_{s \neq i} p_{s \rightarrow j(x_j)} \quad \text{TX nodes} \]

\[ p_i \stackrel{j(x_j)}{\rightarrow} E \left[ \exp \left( u f_i (x_i, z_i) \right) \right] | x_j \quad \text{RX nodes} \]

- Two simplifications for wireless scheduling:
  - RX node update: Computational complexity reduced by Gaussian approximation
  - Messaging overhead: Use linearization to replace unicast messages with broadcast messages

- Similar methods used in many approximate BP algorithms for CDMA multiuser detection & non-Gaussian estimation:
  - Caire, Boutros ('02), Guo-Wang ('06), Tanaka-Okada ('05), Neirroti-Saad ('05), Kabashima ('05), Donoho, Maleki, Montanari ('09), Bayati-Montanari ('10), Rangan ('10)
BP Multi-Round Protocol

Interference → Desired link → Interference

RX2 → TX1 → RX1 → TX2

Round 0
- TX vector x1(0)
- Sensitivity D2(0)
- Interference z1(0) and sensitivity D1(0)

Round 1
- TX vector x1(1)
- Sensitivity D2(1)
- Interference z1(1) and sensitivity D1(1)

Data scheduled along TX vector x1

Shivendra S. Panwar
2011 Wireless Symposium @ Virginia Tech
Spectrum Sharing

• How short-range communication can co-exist with macro transmissions?
  – Femtocell vs macrocell
  – Primary vs secondary users in cognitive nets

• How to mitigate cross-tier interference between overlays and underlays?
  – Femtocells are generally not connected directly into the operator’s core network, backhaul signaling for interference coordination is often limited.
  – Over-the-air signaling may also be difficult due to the power disparities between the femtos and macros, and the large number of femtos per macro.
Overlays and Communication “Under the Radar”

- The solution is to transmit “Under the radar”
  - Preliminary results suggest that IC can work around this problem by having short-range links cancel signals from the macro transmitters and transmitting “under the radar” of the macrocells
  - Such communication avoids any explicit cross-tier coordination
  - Minimal reduction in macrocell rate (4 to 7%)
  - Significant data rate for short-range links

How does it Work?

• Both the uplink (UL) and downlink (DL) bands of the macrocell overlay network are first partitioned into subbands.
  – Using standard fractional frequency reuse (FFR) methods, the macro transmit power is varied across the subbands to create different reuse patterns and interference conditions in each subband.

• Subband partitioning can also create improved opportunities for femtocell communication.
  – Most femto links will be able to transmit in one or more subbands with minimal interference into the macro network.
  – With a load-spillage power control method, femto transmitters can identify the interference effect in each subband and optimize their power distribution across the subbands appropriately.
An Example

Typical locations of macro UEs (MUE1–4) in each subband, along with possible locations for short-range links (SR1–7). SR1–3 can operate in either the UL or DL and do not require interference cancelation (IC). SR4 and 5 would operate in the DL with IC of the DL signal from the macrocell base stations. SR6 and 7 operate in the UL with IC of the macro UE transmissions.
How spectrum is allocated

• Interference coordination
  – MUE1 and MUE3 use f0 since they are close to the BS;
  – MUE2 and MUE4 use f1 and f2, respectively;
  – SR1 use f0 since it is at the macrocell edge;
  – SR2 or SR3 is close to the BS, they can use a band not used by the BS.

• Interference cancellation
  – SR4 being close to macrocell 1, will receive the DL signal on f1 strongly. Also, the signal on that subband will typically be at low rate and can be easily decoded to perform IC. Same situation for SR5.
  – SR6 and SR7 are subject to strong uplink interference. They perform IC for uplink.
Interference and the Backhaul

• In het nets, interference conditions are closely tied to the backhaul
  – Cell selection decisions depend not only on airlink factors but access restrictions, backhaul limitations, etc.
  – Backhaul conditions also determine level of backhaul-based cooperation / interference coordination

• Need for cross-layer approach to cell selection and interference coordination
  – A new interface between the backhaul and wireless devices
  – Joint optimization of interference coordination / cell selection
Heterogeneous and Mobility

- From Gateway Architecture to Intelligent Edge
What are the Benefits?

• EPC is motivated by several considerations:
  – It offers a constant IP point of attachment that makes the mobility of users 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.
  – Simplifies the design of terminal with centralized processing, with the terminals only providing measurement reports.

• Why switch to Intelligent Edge?
  – Scalability
  – EPC handover have been designed generally assuming that the air–interface is the bottleneck and rely on extensive, fast, messaging in the core network.
  – May perform poorly in a heterogeneous networks, where access points could have limited backhaul or significant delay to the core network.

• Cellular standards have already begun considering support for providing an IP point of attachment directly at the base stations: Selected IP traffic offload (SIPTO) and local IP access (LIPA).

• If mobility is support for those protocols, edge networking can provide the basis for a more scalable architecture without a costly dedicated core network, and enable flexible integration of much 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 (MBB) handover by supporting simultaneous physical-layer connections to both source and target base stations.

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

- During fast handovers, it may be too time-consuming to switch the IP point of attachment.
- Over the air, PHY-layer relaying may be more appropriate.
- Multiple streams of data is forwarded to the relay e-Node B to assist transmissions or re-transmissions.
Cooperation in Heterogeneous Wireless Networks

• Peak data rate has increased dramatically over the years, but problems still exist at 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 station at the edge takes much longer time
  – 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 improved because edge nodes transmit at higher rate thus improving spectral efficiency
  – Candidate relays?
    Mobile user, macro/pico-cell BS, fixed relays, femtocell BS, etc.
  – What are the incentives? Throughput, power, interference
Cooperative / multihop communications have been adopted in the 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 into 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 the macrocell based deployment, the number of operator deployed relay stations is limited
  - In traditional networks, performance gain for cooperation is limited unless user (MS) cooperation is enabled
  - User cooperation gives rise to the following problems: battery consumption, synchronization, security and incentive

- The proliferation of pico/femto base stations will provide enough relays
  - They do not have battery consumption problem
  - They are easier to synchronize:
    - stationary, backbone connection and better radio design
  - They are more secure because they are part of the operator’s network
Cooperative MIMO for Het Nets

- For high mobility MSs or MSs that have not been covered by any femtocell, cooperative MIMO
  - enables fully *opportunistic* use of all available surrounding radios.
  - increases network capacity and helps to reduce coverage holes.
Limitations of previous cooperative methods:
- Single relay: low spatial diversity gain
- Multiple relays: consume more bandwidth resource when several relays sequentially forward signal

Any alternative?
- Distributed Space-Time Coding (DSTC)
- How does DSTC work?
  - Recruit multiple relays to form a virtual MIMO
  - Each relay emulates an indexed antenna
  - Each relay transmits encoded signal corresponding to its antenna index
- Pros: Spatial diversity gains
- Cons:
  - Tight synchronization required
  - Relays need 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
  - randomized 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 greatly reduced
  - Performance comparable to centralized MIMO is attained
R-DSTC for Diversity

- Randomized Distributed Space-Time Coding (R-DSTC)
- How does R-DSTC work in PHY?
  - Two-hop network: source station, relays, destination station.
  - Relays re-encode the first-hop signals and forward over the second hop
    - Unlike DSTC, R-DSTC relay does NOT transmit the signal from a specific indexed antenna
    - Instead, each relay transmits a weighted linear combination of all streams of an underlying STC codeword of size $L \times K$.
  - As long as the number of relays $N > L-1$, a diversity order of $L$ is achieved.
### R-DSTC Advantages

#### Performance Comparison

<table>
<thead>
<tr>
<th></th>
<th>DSTC</th>
<th>R-DSTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only selected relays forward. Low diversity gain.</td>
<td>All relays that overhear hop-1 signals can relay. High diversity gain.</td>
<td></td>
</tr>
<tr>
<td>Global and latest channel information REQUIRED for rate selection.</td>
<td>Detailed channel information NOT REQUIRED; outdated estimates can be used.</td>
<td></td>
</tr>
<tr>
<td>STC codeword allocation REQUIRED.</td>
<td>STC codeword allocation NOT REQUIRED; transmissions can simply be randomized.</td>
<td></td>
</tr>
<tr>
<td>Tight synchronization among relays REQUIRED.</td>
<td>Tight synchronization among relays NOT REQUIRED.</td>
<td></td>
</tr>
<tr>
<td>Received power unbalanced.</td>
<td>Average received power from all relays balanced.</td>
<td></td>
</tr>
<tr>
<td>Performance degrades whenever any selected relay fails to relay.</td>
<td>Full diversity order of L is reached when N&gt; L-1.</td>
<td></td>
</tr>
</tbody>
</table>
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 1/2, 2/3, 3/4
- 20 MHz bandwidth
- Contention window: 15 - 1023
- Transmit power: 100mW

![Graph showing throughput and delay with number of subscriber stations]


Shivendra S. Panwar 2011 Wireless Symposium @ Virginia Tech
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.

Source: C Nie, P Liu, T Korakis, E Erkip and S Panwar, CoopMAX: A Cooperative MAC with Randomized Distributed Space-Time Coding for an IEEE 802.16 Network, IEEE ICC 2009

Shivendra S. Panwar 2011 Wireless Symposium @ Virginia Tech
R-DSM for Spatial Multiplexing

- Mismatch in the number of antennas on BS and MS
  - Assuming each mobile station has only one antenna and the base station has L antennas

- Randomized Distributed Spatial Multiplexing (R-DSM) is based BLAST scheme

- The channel capacity between the relays and the destinations scales linearly with min(N,L), where N is the number of relays

- How does R-DSM work in PHY?
  - Two-hop network: SISO transmission from source to relays first, followed by relays transmitting together to the destination using R-DSM.
  - Each relay independently generates a random coefficient and then transmits a weighted sum of the signals for each antenna in BLAST scheme

Shivendra S. Panwar

2011 Wireless Symposium @ Virginia Tech
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

Source: P Liu and S Panwar, Randomized Spatial Multiplexing for Distributed Cooperative Communications, IEEE WCNC 2009
Cooperative Video Multicast

• Performance of conventional video multicast schemes in an access network is limited

- Source transmits at the lowest transmission rate
- Receivers with good channel quality unnecessarily suffer
Cooperative Video Multicast with R-DSTC

- Source station transmits a packet
- Nodes who receive the packets become relays which re-encode the first-hop signals and forward over the second hop
- Each relay transmits a weighted linear combination of all streams of an underlying STC with a dimension of L
Results: Single Layer Schemes

Cooperative Handover for Pico/Femtocells

- Handover happens much more frequently for MSs in Heterogeneous Networks
  - Smaller BS coverage area
  - Low transmission power of BS
  - Higher signaling overheads and more dropped calls
  - One solution is to let MSs with moderate to high mobility connect to macrocell BS, which reduces overall spectral efficiency

- Cooperative handover in Heterogeneous Networks
  - Separate signaling and data paths
    - Macrocell BS orchestrates handover and allocates radio resources for data transmissions
    - User data goes through surrounding pico/femtocell BSs
Cooperative Handover in detail

- Macrocell BS tracks the locations of the MS and makes predictions for the events of handover and which pico/femtocell BSs the MS is moving to.

- In the downlink
  - Macrocell BS pre-fetches user data packets to a cluster of pico/femtocell BSs via their backhauls
  - Macrocell BS allocates frequency/time slots for the downlink data and broadcasts to all BSs under it
  - Pico/femtocell BSs cooperatively transmit to the MS using R-DSTC

- In the uplink
  - Macrocell BS broadcasts the allocated frequency/time slots for the MSs
  - Any pico/femtocell BS that successfully decodes an uplink user packet forwards it to the Macrocell BS via its backhaul
Coordinated Multipoint Transmission (CoMP)

• Another scheme to improve spectral efficiency at cell edge

• Considered by 3GPP LTE-Advanced as a tool to improve coverage, cell-edge throughput, and/or system efficiency

• By coordinating and combining signals from multiple antennas, mobile users can enjoy consistent performance whether they are close to the center of an LTE cell or at its outer edges

Shivendra S. Panwar

2011 Wireless Symposium @ Virginia Tech
Downlink CoMP

• Joint processing
  – Data is available at each point
  – Joint transmission and fast cell selection (FCS) are being studied
    • Joint transmission: data transmitted from multiple point at a time
    • FCS: data transmitted from one point at a time

• Coordinated scheduling/beamforming (BF)
  – Data is only available at serving cell but user scheduling/BF
  – Decisions are made with coordination among cells

Source: LTE Rel-9 and LTE-Advanced in 3GPP, Takehiro Nakamura, NTT Docomo
Uplink CoMP

- Coordinated multi-point reception
  - Uplink signal is received at multiple points
  - Scheduling decisions can be coordinated among cells to control interference
The Second-Last Mile Problem

• Explosively growing traffic demand
  – More than 5 billion cell phones by 2010
  – Increasing number of data intensive applications
  – 3G/4G standards are pushing up the macrocell data rates (~100 Mbps)

• Poor Cellular infrastructure
  – Most of the BS backhauls use four to six T1/E1 lines (~8 Mbps)
  – Adding BSs or updating data lines is expensive
    (more than $10,000 per line and $50,000 per site annually)

Macrocell backhaul has become the bottleneck!
Solution: FemtoHaul

• System Architecture for FemtoHaul
  – FemtoHaul is a novel solution to the macrocell backhaul problem
  – In FemtoHaul, the femtocell backhaul is used to carry non-femto user traffic by forwarding through a relay

• Detailed Design
  – Channel allocation mechanism based on OFDMA WiMAX
  – Policy for base stations to schedule user transmissions
Simulations demonstrate that our solution can significantly reduce the macrocell backhaul traffic while still guaranteeing a high rate to the subscribers.

Source: A Rath, S Hua and S S Panwar. FemtoHaul: Using Femtocells with Relays to Increase Macrocell Backhaul Bandwidth, IEEE INFOCOM Workshops 2010
The Urge to Merge

• Cellular infrastructure lags behind user demand

• Huge cost to increase capacity – no Moore’s law for wireless!

• What happens when cellular operators merge and pool resources
  – Share infrastructure
  – Share spectrum
How can BS’s be Shared

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

• Sharing the base stations
  – If the users can freely access the BS’s of either operator using the “closest-first” rule
  – Network capacity can be improved
Co-located Cell Towers

• Scenario 1: Base stations from the two operators are co-located on the same cell tower
  – Multiplexing gain by opportunistically connecting to the BS with lower traffic load.
Disjoint Cell Towers

- **Scenario 2:** Base stations from the two operators ("Red" and "Blue", using yellow and light blue spectrum, respectively) are maximally spatially distributed from each other.

- **Diagram notes:**
  - Blue Dashed Line: original cell
  - Blue Solid Line: cell after pooling

- A frequency reuse-2 system
- Halve cell size
- Reduces inter-cell interference
- Reduced SINR

Shivendra S. Panwar 2011 Wireless Symposium @ Virginia Tech
Back-of-the-Envelope Calculation

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

- 3x theoretical gain for edge users in Scenario 2

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

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

IEEE 802.16e
More Realistic Cells

• Similarly in a hexagonal layout
  – Downlink edge user performance
  – Performance of point A, C and E remains the same. Point 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

*Blue Dashed Line*: original cell
*Blue Solid Line*: cell after pooling
Further Improvement

• Without adding more hardware, Scenario 2 triples edge network capacity

• What will happen if new hardware is added to the newly opened cell towers?

• Scenario 3: (Adding more infrastructure)
  – Aggregate the two spectrums at each BS location.
  – 2X gain for aggregate capacity
Cooperation Communications for Battlefields

- Much more reliable communications can be improved between the source and the destination.
  - With only one relay, the error performance is improved.
  - When multiple relays are present, R-DSTC based cooperation scheme is appropriate.
  - Only requires minimum signaling overhead and achieves full diversity.
Immune to Jamming

- Cooperative communications is inherently immune to jamming.
  - The virtual array formed by the cooperative nodes has a large aperture.
  - Allows the remote receiver to use spatial filtering to null the jammer and harvest the power of the useful signals from other directions.
Range Extension

- Transmissions reach much further.
  - Range for a single radio is limited due to its power limit.
  - Distributed space-time codes also help to extend range because of higher reliability.
  - Cooperative beam-forming greatly improves range.
Cooperative Spatial Multiplexing

- Utilize antennas on several nodes to achieve higher diversity gain
  - BLAST gain is limited by $\min(Tx, Rx)$ and the number of antennas on single device is limited
  - Source disseminates its information to a number of relays in the first hop
  - Relays transmit using BLAST coding
  - Capacity limit is $\min(N*Tx, Rx)$
Video Optimization for Wireless

- Video, as the dominate traffic on wireless Internet, can be optimized for the wireless bottleneck

- **Video Transport Optimization**
  - New systems approach to realizing practical cross-layer optimizations
  - Utilize new visual perception metrics in end-to-end transport
  - Multi-user resource allocation in highly congested environments
  - Cooperation models and incentives for het. networks (cell, broadcast, WiFi)
  - Leveraging meta-data / side-information in transport layer

- **Video Processing Optimizations**
  - Video Frame Rate Up-Conversion (FRUC), super-resolution, error concealment
  - Saliency and video analytics (e.g., face detect/tracking for video conferencing) to identify priority regions for unequal error protection

- **Novel Network Architectures**
  - Lower barrier to dramatically increase # network attach points (coop. clients/P2P, caching, heterogeneous cognitive networks, vehicular P2P)
Scalable Video Coding

• A research topic over 20 years
• Prior standards support:
  – H.262/MPEG-2
  – H.263
  – MPEG-4 Visual

• Not successful in prior standard, due to higher costs in:
  – Coding efficiency vs. single-layer codec
  – Decoder complexity

• H.264 Annex G (SVC): 11/2007, final draft released
  – Transport specification issued draft in 2010
SVC Modalities

- Temporal (frame rate) scalability: the motion compensation dependencies are structured so that complete pictures can be dropped from the bitstream.

- Spatial (picture size) scalability: video is coded at multiple spatial resolutions.

- SNR/Quality/Fidelity scalability: video is coded at a single spatial resolution but at different qualities.

- Combined scalability: a combination of the 3 scalability modalities described above.
Perceptual Video Quality

- Compression efficiency improves 2-3x every 10 years
  - Traditionally, codecs are judged by peak signal-to-noise ratio (PSNR)

- H.265 (latest standard in development)
  - Only 20% from objective / “math”
  - Most from visual perception
Challenges

- Evaluating and optimizing the quality of digital imaging systems with respect to the capture, display, storage and transmission of visual information is one of the biggest challenges in the field of image and video processing
  - PSNR doesn’t necessarily corresponding to human assessment on quality
  - Distortions caused by lossy compression methods and transmission errors depends highly on the content
  - Subjective experiments on perceived quality are complex and time-consuming
  - It is necessary to have automatic methods for video quality assessment for various applications, such as video playback on PC’s, video streaming to mobile handsets
Modeling quality, rate, and complexity of scalable video encoding

- Sender codes a video into a scalable stream with fine granularity layers corresponding to different spatial, temporal, and amplitude resolutions (STARs), and order these layers according to their impact on perceptual quality.
  - The maximum STAR depends on the expected (long-term) maximum sustainable video rate between the sender and receivers.

- At the sender and each intermediate node, determine the expected (short-term) maximum sustainable video rate in the remaining links, extract and send only those layers up to the sustainable rate.

- If the receiver/sender is limited in processing energy/display, also consider energy/display constraint
Research Components

• Modeling perceptual quality and rate of video as functions of STAR
• Ordering video layers and packets according to rate and quality models
• Modeling the complexity and memory access of video encoding and decoding, respectively, as functions of STAR
• Using the complexity and memory models to adapt the voltage/
• Frequency/memory allocation during video encoding/decoding to reduce energy consumption
• Model the energy consumption as a function of STAR
• Using the resulting model for video adaptation subject to bandwidth and energy constraint
• Unequal error protection based on quality and rate model
Proposed Models

Rate model

Quality model

Source: QSTAR: A Perceptual Video Quality Model, Yao Wang, NYU-Poly

Shivendra S. Panwar

2011 Wireless Symposium @ Virginia Tech
Millimeter Wave and Sub-THz Wireless Communications

- Recent advances in CMOS processes made millimeter wave bands (30-300 GHz) possible for wireless personal communications

- Global spectrum regulators have agreed upon several GHz of license-free wireless personal area network (PAN) spectrum centered around 60 GHz
US Frequency Allocation Chart
Unleashing the 30-300GHz Spectrum

If 40% of the spectrum in the mmW bands is available, 100+ GHz new spectrum for mobile broadband!

More than 200 times the spectrum currently allocated for this purpose below 3GHz.
Radio Characteristics

• Propagation Loss
  – Free-space loss
  – Atmosphere Gaseous Losses
  – Precipitation attenuation
  – Foliage blockage
  – Scattering
  – Bending

• Propagation and other losses due to rain, foliage and penetration through building materials needs better understanding

• Attenuation higher in most materials

• Millimeter waves are attractive for mobile application due to small component sizes such as antennas
  – More antennas can be integrated
  – Higher MIMO gains
  – Highly directional antennas to avoid interference
Use Cases

• Shorter-range applications
  – 60 GHz, 120 GHz, 183 GHz, 325 GHz, and 380 GHz
  – Personal Area Network (PAN) and Local Area Network (LAN)

• Longer-range applications
  – 100 GHz and 240 GHz
  – Wireless mm-Wave and sub-mm-wave backhaul

• Recent Standardization Efforts
  – WirelessHD: Uncompressed HD video
  – ECMA-387: Bulk data transfer and HD streaming
  – 802.15.3c (TG3c): Point-to-point file transfer & streaming
  – 802.11ad (TGad): File transfer & HD streaming
  – WiGig: File transfer, wireless display & HD streaming
Summary

• The future network architecture is heterogeneous, with macro-, pico- and femto-cells, along with WiFi and (some) ad hoc nodes

• In our view, a large part of the capacity increase predicted will be drained by increased deployment of WiFi, femto/picocells for stationary or slow moving users

• Femtocells, in particular, are the carrier’s Trojan Horses!

• Macrocell bandwidth, which is really the main bottleneck, should be used only when there is no alternative (like satellite networks are today)

• Cooperative networking can be used in such emerging environments by using user end devices, femtocells, WiFi access points, picocells, and macrocell infrastructure as the devices that constitute the cooperating nodes