# Sensor Networks

## Part 5: Sensor Network Signal Processing

**CATT Short Course, March 11, 2005**

Mark Coates

Mike Rabbat

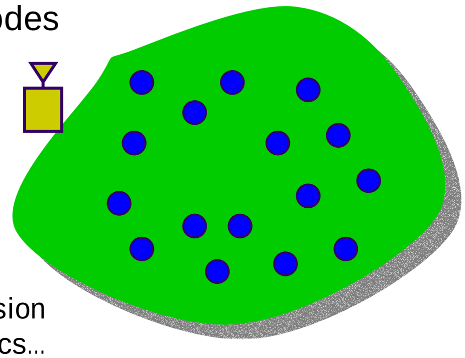


## Collaborative Signal Processing

- "The Network is the Sensor"
  -Oak Ridge Nat'l Labs
- Pottie & Kaiser, 2000
  - Transmit 1Kb over 100m @ 1GHz $\approx$ 3J
  - Equivalent to 3 million processor cycles
- Pister et al., 2001
  - Sense: 4nJ/sample
  - Compute: 1pJ/instruction
  - Communicate: 100nJ/bit
- Culler et al., 2001 (in low power mode)
  - Transmit: 12mW
  - Receive: 5mW
  - Listen: 0.5mW

## Our Mission: Energy Savings

- Network of **n** sensor nodes
- One value each
- Compute the average
  - ML estimate
  - Soft or hard decision fusion
  - Collecting simple statistics…
- Can in-network processing save energy?
- Consider total number of transmissions

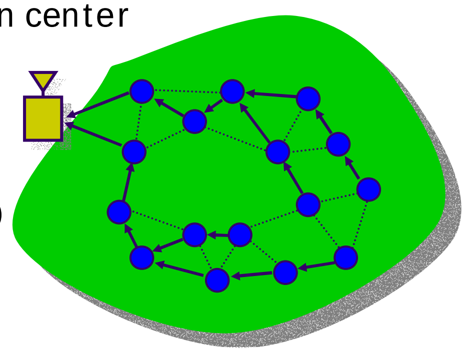


Part 5: 3

## Tree-Based Routing

- Multi-hop data to fusion center
- Relay raw data

$$\mathcal{E}(n) = \sum_{i=1}^{n} (\#i\text{'s children}) \approx n \log n \text{ transmits}$$



- Accumulate at each node (process in-network)

$$\mathcal{E}(n) = n \times 1 \text{ transmits}$$

Part 5: 4

## Consensus

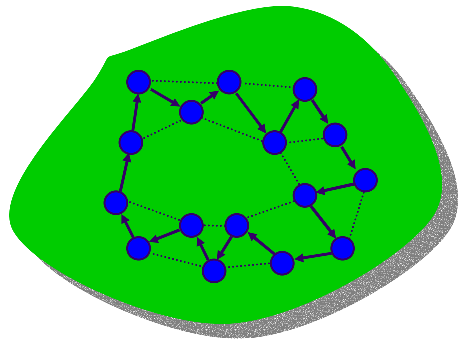
- All nodes get the answer
- Assume Hamilton Cycle
- Relay raw data

$$\mathcal{E}(n) = n(n-1) = O(n^2)$$

- In-network accumulation

$$\mathcal{E}(n) = 2n = O(n)$$

- More generally, exchange with neighbors or groups



Part 5: 5

## Energy Can Be Saved!

- Aggregation

$$O(n \log n) \quad \text{vs.} \quad O(n)$$

- Consensus

$$O(n^2) \quad \text{vs.} \quad O(n)$$

- So, what about in real problems?

Part 5: 6

## Today's Whirlwind Tour

- Decentralized Estimation
- Security Issues
- Source localization
- Tracking
- Distributed Detection
- Field Estimation

Part 5: 7

## Decentralized Estimation

- Simple in-network processing schemes, a.k.a. "data aggregation"
  - Max, min, mean
  - Tree-based schemes (Krishnamachari et al., '02)
- Distributed Maximum Likelihood Estimation
  - (Blatt & Hero, '04)
  - Information theoretic formulation, handles general densities
  - Nodes exchange sufficient statistics with every other node
- Cycle-based algorithms
  - Distributed EM Algorithm for Gaussian Mixtures (Nowak, '03)
  - Incremental Distributed Optimization (Rabbat & Nowak, '04)

Part 5: 8

## Localized Consensus Algorithms

- (Scherber & Papadopoulos, '04), (Xiao & Boyd, '05)
- Measurement Model $y_i = \theta + w_i, \; i = 1, \ldots, n$

$$w_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$$

- MLE given by $\hat{\theta}_{ML} = (1^T y)/n$
- Exchange values with neighbors, $N_i$, at each iteration
- Linear updates at node **i**

$$x_i(t+1) = W_{ii} x_i(t) + \sum_{j \in N_i} W_{ij} x_j(t)$$

- In vector notation, $x(t+1) = Wx(t)$



## Convergence Analysis

- Equivalently, $x(t+1) = Wx(t) = W^t x(0), \; x(0) = y$
- Desire $\lim_{t \to \infty} W^T x(0) = (1^T y/n)1 = (\hat{\theta}_{ML})1$

$$\lim_{t \to \infty} W^T = 11^T/n$$

- Translates to the requirements:

$$1^T W = 1^T, \; W1 = 1, \; \rho(W - 11^T/n) < 1$$

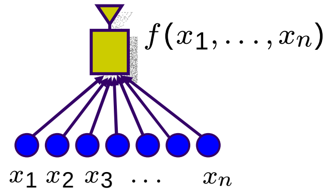
- Rate of convergence

$$\sup_{x(0) \neq \hat{\theta}_{ML} 1} \lim_{t \to \infty} \left( \frac{\|x(t) - \hat{\theta}_{ML} 1\|}{\|x(0) - \hat{\theta}_{ML} 1} \right)^t = \rho(W - 11^T/n)$$

## Resilient Aggregation

(Wagner, '04)

- What if some nodes are malicious?




- Average, sum, product are insecure

$$f(x_1^*, x_2, ..., x_n) = x_1^* + x_2 + ... + x_n$$

- Count, median are more secure

$$f(x_1, ..., x_n) = \sum_{i=1}^{n} \{x_i < \gamma\}$$

Part 5: 11

## Source Localization

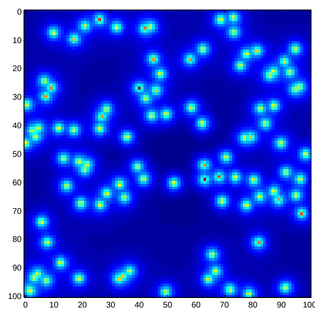
- "A canonical problem in sensor networks"
  - Feng Zhao
- Localize an isotropic energy source
  - Shooter localization
  - Isolating and counting frogs or crickets
- Measurement modalities
  - Time-difference of arrival
  - Direction of arrival
  - Received signal strength

Part 5: 12

## Range-Based Source Loc

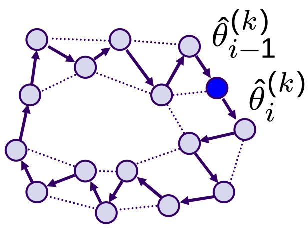
- Typical received signal strength formulation
  - Sensor i at known location $x_i$
  - Measurements $y_i \sim \mathcal{N}\left(\frac{a}{\|x_i - \theta\|^\beta}, \sigma^2\right)$
- Maximum Likelihood Estimate

$$\hat{\theta} = \arg\min_\theta \sum_{i=1}^{n} \left(\frac{a}{\|x_i - \theta\|^\beta} - y_i\right)^2$$




## Distributed Gradient-Descent

- Set $f_i(x_i, y_i, \theta) = \left(\frac{a}{\|x_i - \theta\|^\beta} - y_i\right)^2$ (Rabbat & Nowak, '04)
- Want to compute $\hat{\theta} = \arg\min_\theta \sum_{i=1}^{n} f_i(x_i, y_i, \theta)$
- "Incremental" gradient descent

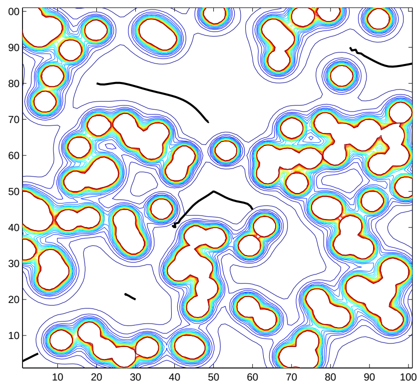



$$\hat{\theta}_i^{(k)} = \hat{\theta}_{i-1}^{(k)} - \mu \nabla f_i(x_i, y_i, \hat{\theta}_{i-1}^{(k)})$$

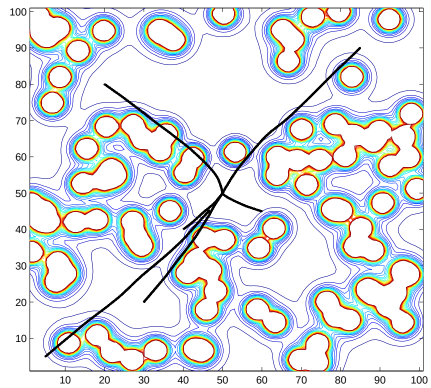
$$\hat{\theta}_n^{(k)} = \hat{\theta}_0^{(k+1)}$$

Part 5: 14

## Simulated Examples




Standard Centralized Method




Incremental Gradient Descent

Part 5: 15

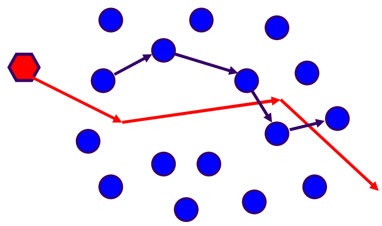
## Tracking

- Obvious military applications
- Generally same types of measurements as for source localization
- Well-studied centralized problem
- Solutions not easily extended to distributed, resource-constrained setting
- Beamforming, Kalman filter extensions

Part 5: 16

## Information-Directed Tracking

(Zhao et al., '03) (Wang et al., '04)



- Choose next sensor to maximize utility

$$\hat{j} = \arg\max_{j \in N_i} \alpha\phi(\underbrace{p(\theta|z_i^t, z_j^t)}_{\text{belief}}) + (1-\alpha)\psi(z_j^t)$$

utility function

cost of using j

Part 5: 17

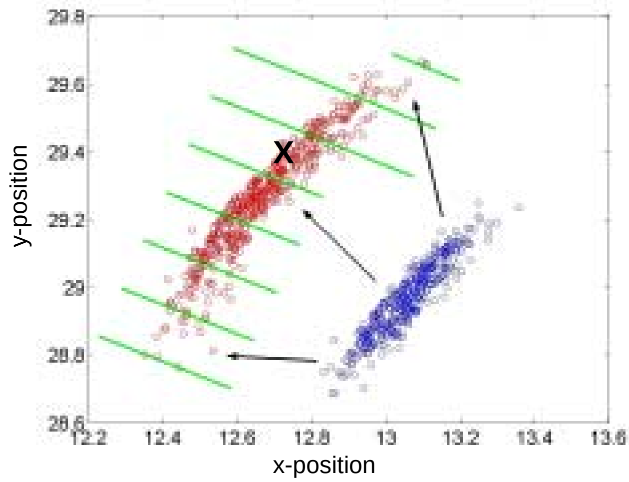
## Distributed Particle Filtering

(Coates, '04)

- Particle Filters
  - Keep track of a collection of plausible states ("particles")
  - Evaluate how well they (1) conform to the dynamic model and (2) agree with the data
- Challenges
  - Computationally intensive
  - Lots of particles to transmit/store
- The Coates Solution
  - Hybrid network setup
  - Maintain common particle filter at multiple nodes
  - Use particle representation for estimation and compression

Part 5: 18

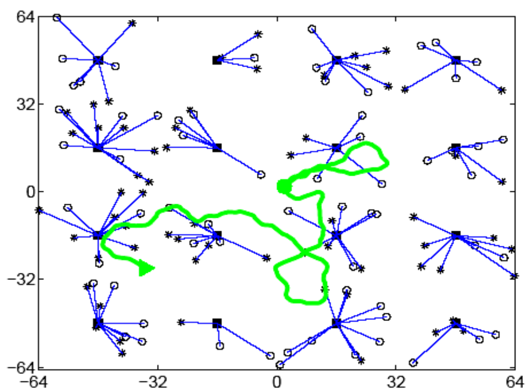
## Compressing New Samples




Part 5: 19

## Simulation Example

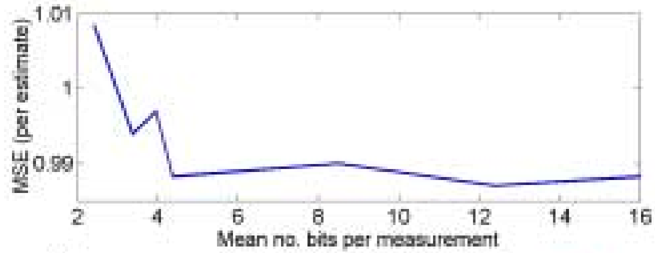
- 16 Class A nodes
- 128 Class B nodes
- 8 Class B nodes measure dist. to object plus noise at each time step
- Object either goes straight, right, or left
- Example of 500 steps



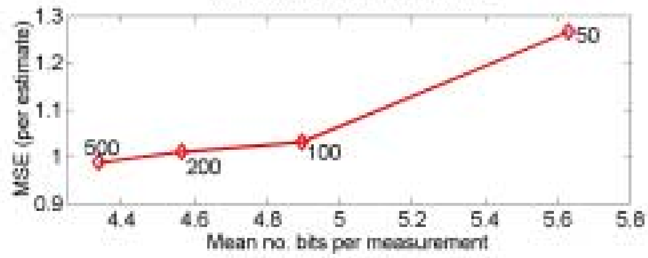
Part 5: 20

# Example Performance

500 particles, varying quantization




32 bit quant, different # particles



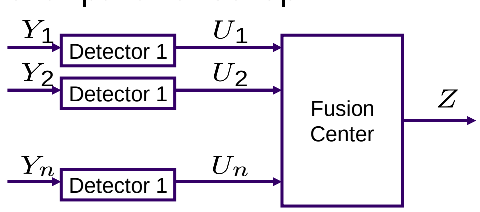

# Distributed Detection

- Binary decision problem

$$\begin{aligned} H_0 &: Y_i \sim p(y|H_0) \\ H_1 &: Y_i \sim p(y|H_1) \end{aligned}$$

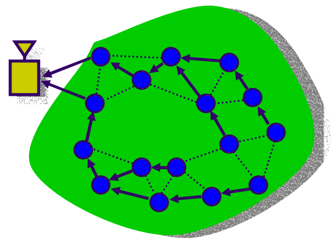
$$\left( \text{E.g.,} \quad \begin{aligned} H_0 &: Y_i \sim \mathcal{N}(0, \sigma^2) \\ H_1 &: Y_i \sim \mathcal{N}(\mu, \sigma^2) \end{aligned} \right)$$

- Traditional parallel setup




- Problem: Doesn't look much like a WSN!

# “Censoring Sensors”

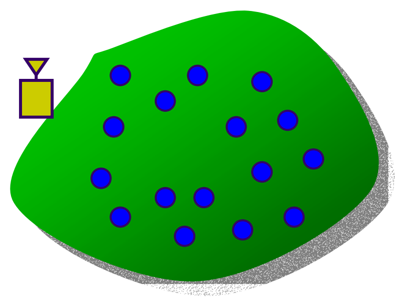


(Blum et al., ’96)
(Appadwedula et al., ’02)
(Patwari & Hero, ’04)

- Transmit to parent when LR > threshold
- Silence is implicitly observed
- Constrain expected # transmissions, optimize detection probability
- Conclusion: Communicate less, but take a performance hit



# Field Estimation



Goal: Measure and convey a physical field

(Temperature, humidity, pressure, light, chemical concentration, elevation)

## Challenges & Considerations

- Physical limitations
  - Low density network = low resolution
  - High density = high resolution
- Density-related quantities to balance
  - Network capacity
  - Energy consumption
  - Error in reconstructed signal
- Balance by adjusting
  - Which nodes transmit
  - What gets transmitted

Part 5: 25

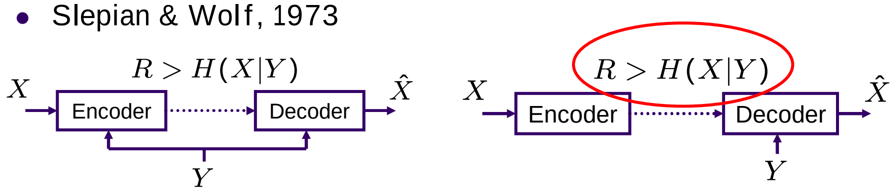
## Two Approaches

1. **Distributed Compression**
   - Sensors sample a stationary random field with mean zero, known covariance
   - Goal: Communicate sample values most efficiently
   - Coding with side-information
   - K. Ramchandran, D. Neuhoff, S. Servetto
2. **Comprestimation**
   - Sensors sample "signal plus white noise", field modeled as smooth or piece-wise smooth
   - Goal: Estimate and communicate the signal efficiently
   - Ideas from multi-scale approximation
   - R. Govindan, R. Nowak, B. Krishnamachari

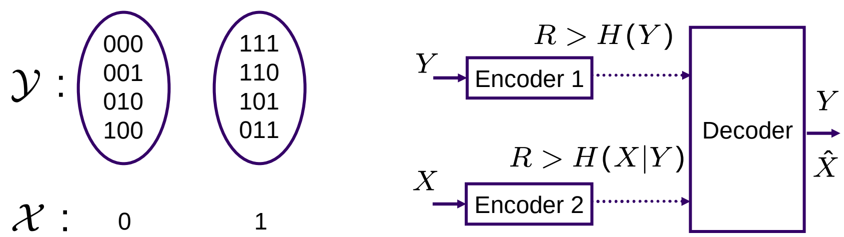
Part 5: 26

# Slepian-Wolf Coding

- Slepian & Wolf, 1973



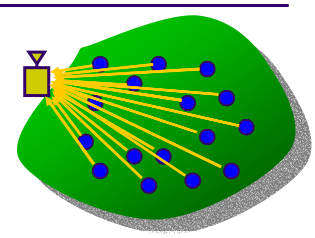

- Pradhan, Kusuma, & Ramchandran, 2002



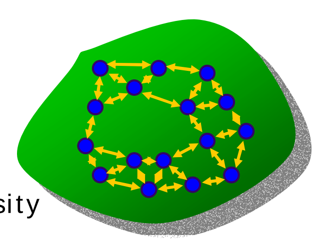



# Distributed Compression Results

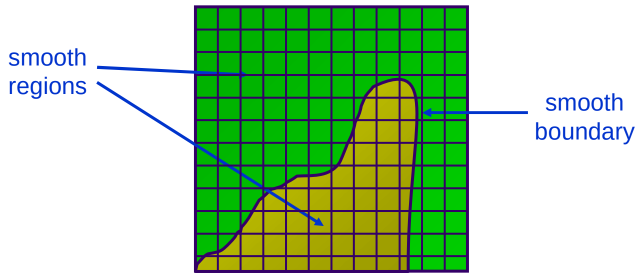
- Many-to-one (Neuhoff)
  - Capacity scales like $O(n^{-1})$
  - Higher density ➔ lower capacity
  - Higher density ➔ higher correlation
  - Conclusion: Oversampling is wasteful, let some sensors sleep if necessary



- Node-to-node (Servetto)
  - Iteratively recode eliminate redundancy
  - Per-node error = D/n
  - Bitrate = log(n/D)
  - Conclusion: Achieve error D/n with finite number of bits, regardless of density



## Inhomogeneous Fields

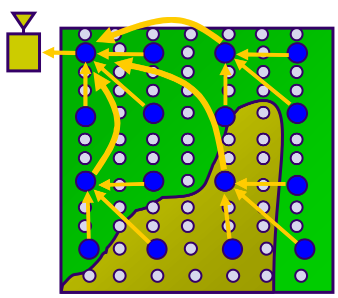



- Tradeoff density, accuracy, & energy
- “Complexity” in estimating smooth boundary

Part 5: 29

## Backcasting

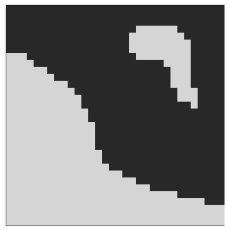
- Hierarchical communication



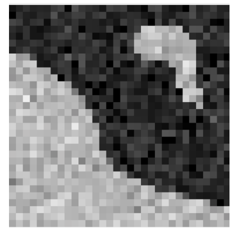
- Stage 1: Coarse estimate (gray nodes sleep)
- Stage 2: Refine interesting regions
- Conclusion: Higher density ➔ Lower MSE, lower energy/node

Part 5: 30

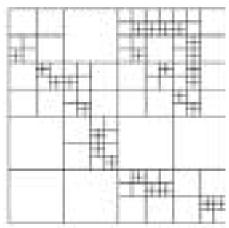
## Backcasting In Action



$\theta^*$



$x$



$\widehat{P}_n$

$\widehat{\theta}_n$



## Summary

- New paradigm for signal processing
  - Distributed, in-network processing
  - Improve efficiency, extend network lifetime
- Plenty of interesting open problems
  - No clear “best solution”
  - No methodical approach
  - Mobility, unreliable networks
  - Communications, beamforming, & virtual arrays, causal inference