Decentralized Parameter Estimation with Wireless Sensor Networks

Signal Processing and Information Theoretic Aspects

C. Antón-Haro and J. Matamoros
{carles.anton,javier.matamoros}@cttc.es

Centre Tecnològic de Telecomunicacions de Catalunya

NEWCOM++ Summer School 2008 - Bressanone (Italy)

Notes

CTTC PhD Fellowship Program 2008

WHO WE ARE

The **Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)** is an **advanced research center** on Information technologies and Communications based in Castelldefels (Barcelona). Current research areas include **Wireless Communications**, **Access Technologies**, **Communication Subsystems**, **Optical Networking** and **IP Technologies** (see http://www.cttc.es for further information).

WE ARE OFFERING...

Four-year PhD fellowships to carry out a PhD thesis within a growing technological center, in a **young and dynamic atmosphere** with possibility to participate in **international R&D projects**. The initial salary is 13.680 €/year and, additionally, the fellowship also covers the fees of the PhD courses to be taken by the candidate. From the second year on and upon completion of the PhD courses, the fellowships turn into regular work contracts with a salary of 15.600 € /year (gross) during the second year and of 17.076 € /year (gross) during the third and fourth years. Every year the gross salary is potentially extensible by a lump sum bonus of up to 1200 €/year.

WE ARE LOOKING FOR...

People with a MS degree in Electrical or Electronic Engineering, Telecommunications, Computer Science or related fields, with a **good academic record**, **strong motivation for research** and interest or knowledge in **topics** such as (but not limited to): packet scheduling and MAC/RRM techniques for wireless communications, cognitive radio, cognitive networks, multi-carrier systems, MIMO communications, channel and network coding, cooperative communications, ultra-wideband, wireless sensor networks, design of microwave devices (filters, oscillators, antennas), positioning systems, signal acquisition/synchronization, management and control in ASON/DWDM optical networks, routing, traffic engineering, QoS and mobility, self-organisation and measurements in cellular and wireless mesh networks.

Applications must be received **on or before July 13, 2008**, including the following documents: (1) **curriculum vitae**, indicating research preferences as well as the title of the Master Thesis, and (2) scanned copy of the **complete academic record** (mandatory), using the web form at: **http://www.cttc.es/PhDCandidates**

Notes

CTTC PhD Fellowship Program 2008

WHO WE ARE

The **Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)** is an **advanced research center** on Information technologies and Communications based in Castelldefels (Barcelona). Current research areas include **Wireless Communications**, **Access Technologies**, **Communication Subsystems**, **Optical Networking** and **IP Technologies** (see http://www.cttc.es for further information).

WE ARE OFFERING...

$\boxed{\textbf{Four-year PhD fellowships}}$ to carry out a PhD thesis ~~within a growing technological~~ center, in a **young and dynamic atmosphere** with possibility to participate in $\boxed{\textbf{international R\&D projects}}$. The initial salary is 13.680 €/year and, additionally, the fellowship also covers the fees of the PhD courses to be taken by the candidate. From the second year on and upon completion of the PhD courses, the $\boxed{\text{fellowships turn into regular work contracts}}$ with a salary of 15.600 € /year (gross) during the second year and of 17.076 € /year (gross) during the third and fourth years. Every year the gross salary is potentially extensible by a $\boxed{\text{lump sum bonus}}$ of up to 1200 €/year.

WE ARE LOOKING FOR...

People with a MS degree in Electrical or Electronic Engineering, Telecommunications, Computer Science or related fields, with a $\boxed{\textbf{good academic record}}$, **strong motivation for research** and interest or knowledge in **topics** such as (but not limited to): packet scheduling and MAC/RRM techniques for wireless communications, cognitive radio, cognitive networks, multi-carrier systems, MIMO communications, channel and network coding, cooperative communications, ultra-wideband, wireless sensor networks, design of microwave devices (filters, oscillators, antennas), positioning systems, signal acquisition/synchronization, management and control in ASON/DWDM optical networks, routing, traffic engineering, QoS and mobility, self-organisation and measurements in cellular and wireless mesh networks.

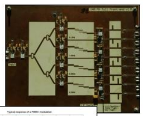
Applications must be received $\boxed{\textbf{on or before July 13, 2008}}$, including the following documents: (1) **curriculum vitae**, indicating research preferences as well as the title of the Master Thesis, and (2) scanned copy of the **complete academic record** (mandatory), using the web form at $\boxed{\textbf{http://www.cttc.es/PhDCandidates}}$

Notes

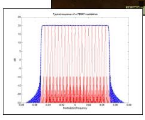
An Image is Worth...

MORE INFORMATION:

http://www.cttc.es/PhDCandidates



http://www.cttc.es/rrhh



Notes

Motivation

This seminar addresses the following issues:

- How to carry out *decentralized* parameter estimation via WSNs.
- To define new network *performance criteria*: estimation accuracy. How accurate can an estimate be?
- Impact of *WSN-related constraints* on performance and network design of: energy consumption and computational complexity constraints, decentralization, scalability properties, etc.
- How to exploit channel fading and and *opportunistic scheduling* to save energy/improve estimation?
- Impact of *quantized observations* on performance?

Notes

References

The sources of this seminar are, mainly:

- Wireless Sensor Networks. Signal Processing and Communications Perspectives. A. Swami, Q. Zhao, Y-W. Hong, L. Tong Eds., John Wiley & Sons Ltd., 2007.
- Fundamentals of Statistical Signal Processing: Estimation Theory. S. M. Kay, Prentice Hall Signal Processing Series, 1993.
- Elements of Information Theory. T. M. Cover and J. A. Thomas, Wiley Series in Telecommunications, 1993.
- Fundamentals of Wireless Communications, D. Tse, P.Wiswanath, Cambridge University Press, 2005.
- Several articles by Slepian & Wolf, Ribeiro & Giannakis, Gesbert & Alouini, Goldsmith, etc.

Notes

Outline I

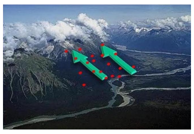
Notes

Outline II

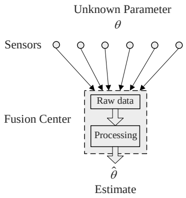
Notes

What is a WSN?

A network with a number of resource-constrained nodes, densely (and often randomly) deployed for one specific purpose, such as sensing or detecting a given phenomenon.

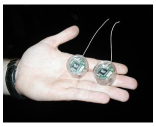


- Typically, measurements are conveyed over wireless channels to a Fusion Center (FC), where data is processed.



Notes

Some practical considerations on WSNs

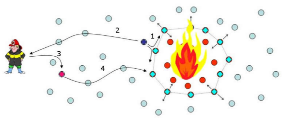


- Sensor nodes with limited sensing, processing and transmission capabilities.
- Sensors are inexpensive devices prone to failures.
- Need for energy-efficient operation
  - A means to maximize network lifetime
  - Sensor nodes are equipped with finite batteries which are difficult/impossible to replace
  - Performance-energy efficiency tradeoffs
- Network size and scalability issues:
  - Do protocols/algorithms/... scale well with the number of nodes?
  - Same software running on all nodes?
- Deployment: massive/random? pre-designed sensor location?

Notes

Typical tasks for WSNs

Detection: To determine whether an event has occurred or not

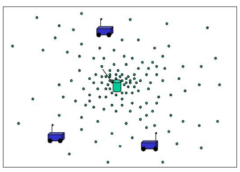


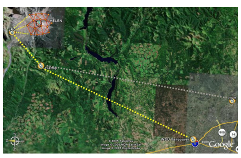
- Applications: earthquake/fire detection, military surveillance,...
- Performance metrics: probability of detection, probability of false alarm, latency, etc.

Notes

Typical tasks for WSNs

Localization/tracking: To determine and/or track the position of goods/people/.. of interest.





- Applications: border control, inventory tracking, logistics and transportation,...
- Performance metrics: positioning accuracy, coverage area, number of simultaneous targets, etc.

Notes

Typical applications of WSNs

Estimation: To accurately represent the observed phenomenon

- Applications: environmental/disaster area monitoring, patient monitoring, pollution measurements,...
- Performance metrics: distortion in the estimates, distortion vs. latency, etc.

Notes

Parameter estimation with WSNs: a good deal?

Pros:

- Increased sensing/monitoring area
- Ability to track spatial variations

Cons:

- Communication with the fusion center/other sensor nodes: fading, path-loss.
- Node reliability: what if 1/+ nodes stop working?

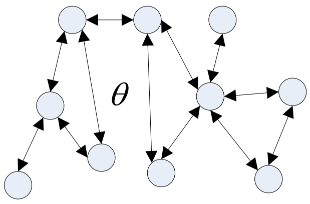
Challenges:

- Need for decentralized estimation techniques - centralized ones to act as a benchmark!
- Availability of local channel state information (CSI) only at the sensor nodes: channel gains, residual energy, etc.
- Energy consumption to acquire channel state information.
- A priori information on the parameter... and how to use it

Notes

Network Topologies

- **Infrastructure-based networks**: A central device (FC) gathers and processes the sensor measurements.
- **Infrastructureless Networks**: Homogeneous network
  - Consensus Networks



Issues: Sensors operate autonomously (no FC coordination), synchronization, belief propagation through P2P communications, stability and convergence time,...

Notes

Background Material: Classical Estimation Theory

Notes

Classical Estimation

Problem Statement:

- Parameter to estimate: θ, deterministic, unknown, scalar (or vector $\boldsymbol{\theta}$)
- Observation vector: $\mathbf{x} = [x_1, \ldots, x_N]^T$ with $\mathbf{x} \in \mathcal{X}$ a $N \times 1$ **random** vector
- Probability density function: $p(\mathbf{x}; \theta)$, known, parameterized by θ.
- Estimator: $\hat{\theta}(\mathbf{x}) = g(\mathbf{x})$

Example: Estimation of DC level in AWGN (σ_n^2)

$$x_k = \theta + n_k \quad k = 1, \ldots, N$$

- Estimator: e.g. : $\hat{\theta}(\mathbf{x}) = \frac{1}{N}\sum_{k=1}^{N} x_k$
- Variance: $\text{Var}\left(\hat{\theta}(\mathbf{x})\right) = \mathbb{E}_x\left[\left(\hat{\theta}(\mathbf{x}) - \mathbb{E}_x\left[\hat{\theta}(\mathbf{x})\right]\right)^2\right] = \frac{\sigma_n^2}{N}$

Notes

Design Criteria For Classical Estimators

Minimum Mean Squared Error (MMSE):

$$\begin{aligned}\text{MSE}\left(\hat{\theta}(\mathbf{x})\right) &= \mathbb{E}_x\left[\left(\hat{\theta}(\mathbf{x}) - \theta\right)^2\right] \\ &= \underbrace{\mathbb{E}_x\left[\left(\hat{\theta}(\mathbf{x}) - \mathbb{E}_x\left[\hat{\theta}(\mathbf{x})\right]\right)^2\right]}_{\text{Var}(\hat{\theta}(\mathbf{x}))} + \underbrace{\left(\mathbb{E}_x\left[\hat{\theta}(\mathbf{x})\right] - \theta\right)^2}_{\text{bias term}}\end{aligned}$$

Intuitive ... but often not feasible !

Minimum Variance Unbiased Estimator (MVU):

- Impose unbiasedness: $\mathbb{E}_x\left[\hat{\theta}(\mathbf{x})\right] = \int_{\mathcal{X}} \hat{\theta}(\mathbf{x})\, p(\mathbf{x}; \theta) = \theta$
- Design estimator with minimum variance:

$$\hat{\theta}(\mathbf{x}) = \arg\min_{\hat{\theta}(\mathbf{x})} \mathbb{E}\left[\left(\hat{\theta}(\mathbf{x}) - \theta\right)^2\right]$$

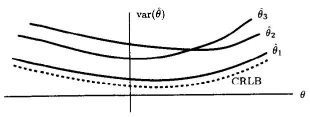
Notes

Cramer-Rao Lower Bound

Definition: Cramer-Rao Lower Bound

An absolute benchmark for any unbiased estimator:

$$CRLB(\theta) = \left(-\mathbb{E}\left[\frac{d^2 \log p(\mathbf{x}; \theta)}{d\theta^2}\right]\right)^{-1} \leq var\left(\hat{\theta}(\mathbf{x})\right)$$



Definition: Efficient Estimator

If it attains the CRLB... for all θ !

Notes

Cramer-Rao Lower Bound

Example: Estimation of DC level in AWGN noise

- Estimator: $\hat{\theta}(\mathbf{x}) = \frac{1}{N}\sum_{k=1}^{N} x_k$
- PDF: $p(\mathbf{x}; \theta) = \frac{1}{(2\pi\sigma_n^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma_n^2}\sum_{k=1}^{N}(x_k - \theta)^2\right]$
- $CRLB(\theta) = \sigma_n^2/N$
- $\text{Var}\left(\hat{\theta}(\mathbf{x})\right) = \sigma_n^2/N$

... so this estimator is the best we can do (MVU)!!

If the minimum variance estimator exists, it attains the CRLB!

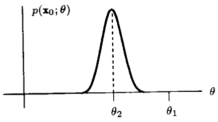
Notes

Some problems associated to MVU estimation

Problems:

- Maximum Likelihood Estimator (MLE): Straightforward & popular

$$\hat{\theta}_{\mathrm{ML}}(\mathbf{x}) = \arg\max_{\theta} p(\mathbf{x};\theta)$$



- Best Linear Unbiased Estimator (BLUE): For linear/linearized data models

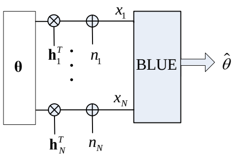
Notes

Best Linear Unbiased Estimator (BLUE)

For linear/linearized vector $(p \times 1)$ observations

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{n} \qquad \mathbf{n} \sim \mathcal{CN}(0, \mathbf{C}_n)$$

$$\mathbf{x} = \begin{bmatrix} \mathbf{h}_1^T \\ \vdots \\ \mathbf{h}_N^T \end{bmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_p \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_N \end{bmatrix}$$



BLUE estimator is linear in the observations and given by:

$$\hat{\theta}_{\mathrm{BLUE}} = \left(\mathbf{H}^T\mathbf{C}_n^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{C}_n^{-1}\mathbf{x}$$

- Only requires first & second order moments of the observation
- Same as MVU if linear data & gaussian noise

Notes

Best Linear Unbiased Estimator (BLUE)

Particular case: Scalar parameter (i.e. $p = 1$) in uncorrelated noise

- Scalar: $\theta \to \theta$, $\mathbf{H} \to \mathbf{h}$
- Uncorrelated noise: $\mathbf{C}_n = \begin{bmatrix} \sigma_{n_1}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{n_N}^2 \end{bmatrix}$
- BLUE estimator:

$$\hat{\theta}_{\mathrm{BLUE}} = \left(\mathbf{h}^T\mathbf{C}_n^{-1}\mathbf{h}\right)^{-1}\mathbf{h}^T\mathbf{C}_n^{-1}\mathbf{x} = \left(\sum_{k=1}^{N}\frac{|h_k|^2}{\sigma_{n_k}^2}\right)^{-1}\left(\sum_{k=1}^{N}\frac{h_k x}{\sigma_{n_k}^2}\right)$$

- with distortion (variance) given by: $\mathrm{Var}(\hat{\theta}) = \left(\sum_{k=1}^{N}\frac{|h_k|^2}{\sigma_{n_k}^2}\right)^{-1}$
- In case of identical noise variances $(\sigma_{n_k}^2 = \sigma_n^2)$: $\mathrm{Var}(\hat{\theta}) = \frac{\sigma_n^2}{N}$

Notes

Background materials:
Bayesian Estimation

Notes

Bayesian Estimation

Problem Statement:

- Parameter to estimate: θ, random, unknown, scalar (or vector $\boldsymbol{\theta}$)
- Observation vector: $\mathbf{x} = [x_1, \dots, x_N]^T$ with $\mathbf{x} \in \mathcal{X}$ a $N \times 1$ **random** vector
- Probability density function: $p(\mathbf{x}, \theta)$, known, joint distribution of $\mathbf{x}$ and θ.

$$p(\mathbf{x}, \theta) = \underbrace{p(\theta)}_{\text{prior pdf}} \; p(\mathbf{x}|\theta)$$

Pros & cons

- Additional (prior) knowledge to improve estimation
- Risk of mismatched/unknown pdf ... CRITICAL!

Design Criteria:

- MAP(Maximum a Posteriori), MMSE (Minimum MSE)

Notes

Bayesian MMSE Estimator

Bayesian Estimation

- Mean Squared Error (MSE)

$$\begin{aligned} \text{MSE}\left(\hat{\theta}(\mathbf{x})\right) &= \mathbb{E}_{\mathbf{x},\theta}\left[\left(\hat{\theta}(\mathbf{x}) - \theta\right)^2\right] \\ &= \int_{\mathbf{x}}\int_{\theta}\left(\hat{\theta}(\mathbf{x}) - \theta\right)^2 p(\mathbf{x}, \theta) d\mathbf{x} d\theta \end{aligned} \qquad (1)$$

- Bayesian Minimum Mean Squared Error (BMMSE) estimator: The *posterior* mean

$$\hat{\theta}_{\text{BMMSE}}(\mathbf{x}) = \mathbb{E}_{\theta|\mathbf{x}}[\theta]$$

with

$$p(\theta|\mathbf{x}) = \frac{p(\mathbf{x}|\theta)p(\theta)}{\int_\theta p(\mathbf{x}, \theta)d\theta} = \frac{p(\mathbf{x}|\theta)p(\theta)}{p(\mathbf{x})}$$

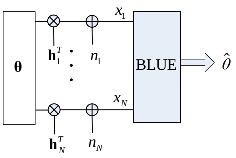
Notes

Bayesian MMSE estimator

Bayesian Estimation

For linear (or linearized) observations

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{n}$$
$$\mathbf{n} \sim \mathcal{CN}(0, \mathbf{C}_\mathbf{n})$$
$$\boldsymbol{\theta} \sim \mathcal{CN}(0, \mathbf{C}_{\boldsymbol{\theta}})$$



The Bayesian (linear) MMSE estimator is:

$$\hat{\boldsymbol{\theta}}_{\text{BMMSE}}(x) = \left(\mathbf{H}^T\mathbf{C}_\mathbf{n}^{-1}\mathbf{H} + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1}\mathbf{H}^T\mathbf{C}_\mathbf{n}^{-1}\mathbf{x}$$

with covariance (distortion) matrix given by,

$$\mathbf{C}_{\theta|\mathbf{x}} = \left(\mathbf{H}^T\mathbf{C}_\mathbf{n}^{-1}\mathbf{H} + \mathbf{C}_{\boldsymbol{\theta}}^{-1}\right)^{-1}$$

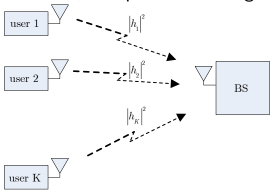
Notes

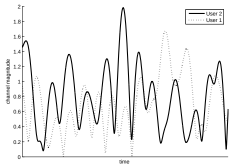
Background Materials:
Multi-user diversity and Opportunistic Scheduling

Notes

Multi-user Diversity and Opportunistic Scheduling

Scenario: Wireless data network, multiple-access (uplink) channel, K users, independent fading for user-BS channels ($|h_1|^2, \ldots, |h_K|^2$)



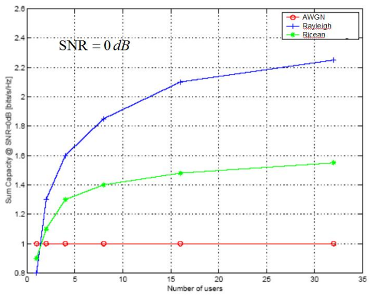


Performance metric: Sum-rate

- Round-robin scheduling: $r_{t,\text{RR}} = \log_2\left(1 + \frac{P|h_{k(t)}|^2}{\sigma_w^2}\right)$
- Opportunistic Scheduling: $r_{t,\text{opp}} = \log_2\left(1 + \frac{P\max_{k(t)}|h_{k(t)}|^2}{\sigma_w^2}\right)$

Notes

Multi-user Diversity and Opportunistic Scheduling



- Multi-user diversity: $\mathbb{E}[r_{t,\text{opp}}] \sim \log\log K$ (Rayleigh Fading)

Notes

Selective Multi-user Diversity

Problem: large amount of feedback (from all users)

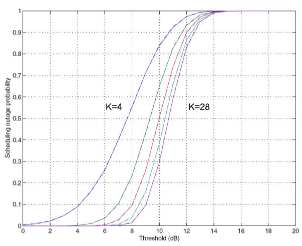
Solution: Selective Multi-user Diversity (SMUD) [Gesbert-Alouini'04]

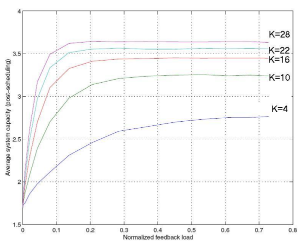
- Algorithm:
  1. Only a subset of users feedback their channel gains: $|h_k|^2 \geq \gamma_k$
  2. BS schedule the user with highest gain out of those above the threshold
- Scheduling Outage: No feedback from sensors - e.g. random selection

$$\Pr(\text{Outage}) = \Pr\left(|h_{\max}|^2 \leq \gamma_{\text{th}}\right)$$

Notes

Selective Multi-user Diversity (cont'd)





- Average SNR = 5dB
- Interplay threshold ⇔ feedback load ⇔ performance
- For $K \geq 25$ users, 10% of users reporting SNRs suffices.

Notes

Multi-user Diversity with Quantized Feedback
Multi-user Diversity and Opportunistic Scheduling

Problem: channel gains feedback in "analog" form
Solution: use quantized channel gains [Nosratinia'07]

- Algorithm:
  1. Users: just notify if above/below threshold $1\{|h_k|^2 \geq \gamma_{\text{th}}\}$
  2. BS: select at random out of those above the threshold
- Performance: Sum-rate
  - If k users report:

$$\bar{r}_k = \sum_{i=1}^{k} \Pr\left[\text{ the } i^{\text{th}} \text{ best user is selected}\right] r_i = \frac{1}{k}\sum_{i=1}^{k} r_i$$

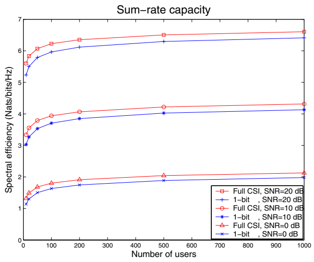
  - For any number of users ($k = 1 \dots K$):

$$\bar{r} = \sum_{k=1}^{n} p_k \bar{r}_k$$

  - Need for optimized choice of the threshold (γ_{th})

Notes

Multi-user Diversity with Quantized Feedback
Multi-user Diversity and Opportunistic Scheduling



- Same growth-rate as with "analog" feedback ($K \to \infty$)
- Small penalty wrt "analog" feedback

Notes

Opportunistic Scheduling and WSNs
Multi-user Diversity and Opportunistic Scheduling

- Differences: The goal now is to maximize estimation accuracy (rather than sum-rate)
- Similarities: To opportunistically schedule sensors results into enhanced performance

Example: Two sources of "randomness"

- Observation noise: $x_k = \theta + n_k$; $n_k \sim \mathcal{CN}(0, \sigma_{n_k}^2)$
- Sensor-FC channels: $|h_k|^2$

Intuition: Distortion minimized for sensors with high $|h_k|^2$ and low observation noise.

Notes

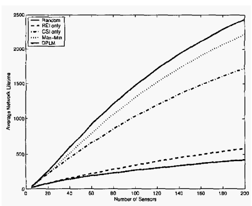
Opportunistic Scheduling & Network Lifetime
Multi-user Diversity and Opportunistic Scheduling

- Extending Network Lifetime (LT) is crucial in WSNs
- Definitions: time elapsed until first sensor dies/part of the WSN is not reachable/estimation quality cannot be achieved/...
- Scheduling with Residual Energy Information (REI)[Chen-Zhao'05]
  - Let $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ be the residual energies.
  - New scheduling rule:

REI & CSI $\quad i^* = \arg\max_i \dfrac{\varepsilon_i}{E_{\text{packet}}(|h_i|^2)}$

CSI only $\quad i^* = \arg\max_i |h_i|^2$

REI only $\quad i^* = \arg\max_i \varepsilon_i$



Combined use of REI & CSI maximizes lifetime (1st sensor dies)

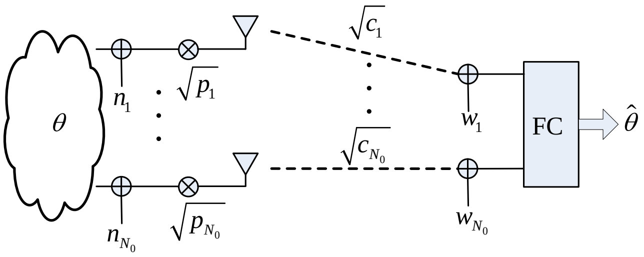
Notes

Notes

Decentralized Estimation with Analog Transmission

Notes

Consider a WSN composed of one Fusion Center (FC) and a large population of N_0 energy-constrained sensors aimed at estimating a scalar, slowly-varying and spatially-homogeneous parameter θ.



Notes

- Each sensor observes $x_i = \theta + n_i$; $i = 1, \ldots, N_0$
- Then, it re-transmits a scaled version to the FC:

$$\mathbf{y} = \mathbf{h}\theta + \mathbf{z} \qquad \begin{array}{l} \mathbf{h} = \left[\sqrt{p_1 c_1}, \ldots, \sqrt{p_{N_0} c_{N_0}}\right]^T \; ; \; c_i = |h_i|^2 \\ \mathbf{z} \sim \mathcal{CN}(0, \mathbf{C_z}) \\ \text{diag}\left[\mathbf{C_z}\right] = \left[p_1 c_1 \sigma_n^2 + \sigma_w^2, \ldots, p_{N_0} c_{N_0} \sigma_n^2 + \sigma_w^2\right] \end{array}$$

- The FC uses a Best Linear Unbiased Estimator (BLUE) to reconstruct θ:

$$\hat{\theta} = \left(\mathbf{h}^T \mathbf{C_z}^{-1} \mathbf{h}\right)^{-1} \mathbf{h}^T \mathbf{C_z}^{-1} \mathbf{y}$$

with distortion given by:

$$D = \text{Var}(\hat{\theta}) = \left(\sum_{i=1}^{N_0} \frac{p_i c_i}{p_i c_i \sigma_n^2 + \sigma_w^2}\right)^{-1}$$

Notes

- Minimization of Distortion:

$$\min_{p_1, \ldots, p_{N_0}} D \qquad \text{s.t.} \qquad \sum_{i=1}^{N_0} p_i \le P_t$$

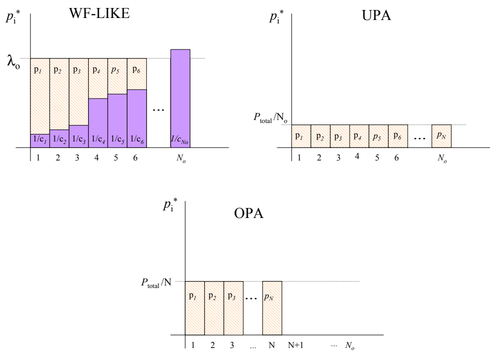
- Solution: Waterfilling-like [Goldsmith et al]

$$p_i^* = \frac{\sigma_w^2}{\sigma_n^2 c_i}\left[\frac{\sqrt{c_i}}{\sqrt{\lambda_0}\sigma_w} - 1\right]^+ \qquad i = 1 \ldots, N_0$$

- Drawback: Extensive signalling sensors $\leftrightarrow$ FC required
- Fix: Opportunistic Power Allocation (OPA) [Matamoros & Antón-Haro]
  - Let only a subset of sensors $N \le N_0$ participate on average in the estimation process and keep FC-sensor signalling low
  - Performance loss??

Notes

Comparison of Power Allocation Schemes
Decentralized Estimation with Analog Transmission

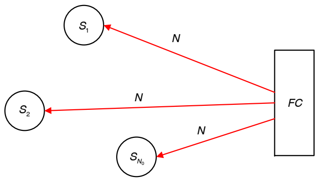


- OPA vs. WF-like and Uniform Power Allocation (UPA)

Notes

Opportunistic Power Allocation (OPA)
Decentralized Estimation with Analog Transmission

1. **Initialization**: Compute the channel gain threshold $\gamma_{\rm th}$.
2. **Identification of the active sensor set**: In each time-slot, each sensor node notifies the FC whether it will participate in the estimation process or not.



3. **Power Allocation and Transmission**: The N active sensor nodes send their observations to the FC.
4. **Go to Step 2.**

Notes

OPA - Threshold Computation
Decentralized Estimation with Analog Transmission

- **Power Allocation Rule**: Uniformly allocated to the subset of **active** sensors, that is,

$$p_i = P_t/N; \qquad i = 1 \dots N$$

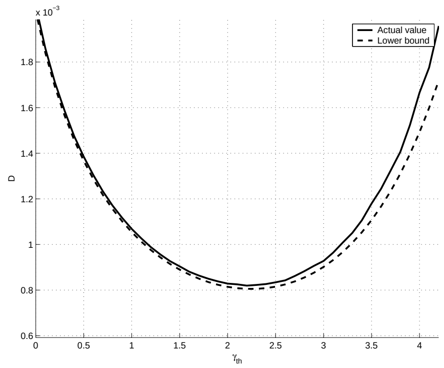
- **Threshold Computation**:
The threshold $\gamma^*_{\rm th}$ which minimizes the bound is given by,

$$\gamma^*_{\rm th} = \left[2\mu_c W_0 \left(\frac{1}{2} \sqrt{\frac{N_0 \sigma_w^2 e}{P_t \sigma_n^2 \mu_c}} \right) - \mu_c \right]^+$$

where $W_0(\cdot)$ is the Lambert function, i.e. the inverse function of $f(W) = We^W$.

Notes

A Sanity Check on the Lower Bound...
Decentralized Estimation with Analog Transmission

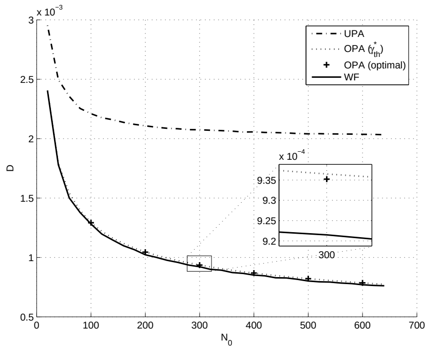


- The lower bound is convex and tight !

Notes

OPA - Performance Results

Decentralized Estimation with Analog Transmission



- OPA outperforms Uniform Power Allocation and very close to WF.
- Almost no penalty for using the approximate threshold γ^*_{th}

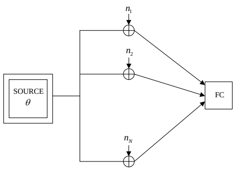
Notes

Decentralized Estimation with Digital Transmission

Notes

Distortion with centralized observations

Decentralized Estimation with Digital Transmission



- θ: deterministic, scalar
- $x_k = \theta + n_k$; $k = 1, \dots, N$
- $n_k \sim \mathcal{N}(0, \sigma_n^2)$

- Sample mean estimator (BLUE & MVU):

$$\hat{\theta}_{\text{BLUE}} = \frac{1}{N} \sum_{k=1}^{N} x_k \quad ; \quad \text{Var}(\theta) = \frac{\sigma_n^2}{N} \tag{2}$$

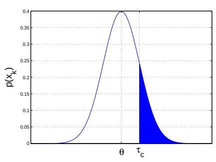
- What can we expect with *decentralized* estimation and *1 bit per sensor* (bandwidth constraints)?

Notes

Quantization with 1 Bit per sensor

Decentralized Estimation with Digital Transmission

- **Sensors**: Send indicator variable $b_k \in \{0, 1\}$... a set of 1-bit quantizers!! [Ribeiro-Giannakis'05]



- b_k: Bernoulli random variable with parameter

$$q_k(\theta) = F(\tau_c - \theta)$$

- F: Complementary CDF of noise
- τ_c: design parameter

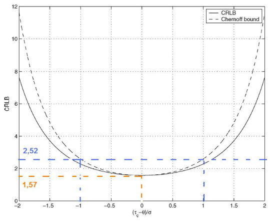
- **FC**: Produces an ML estimate

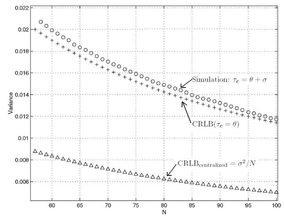
$$\hat{\theta}_{\text{ML}} = \tau_c - F^{-1}\left(\frac{1}{N}\sum_{k=1}^{N} b_k\right)$$

with $\hat{q} = \frac{1}{N}\sum_{k=1}^{N} b_k$.

Notes

- Cramer-Rao Lower Bound: $CRLB(\theta) = \text{Var}(\hat{\theta})$ for large N





- $\text{CRLB}_{\text{min}} \approx 1.57\frac{\sigma^2}{N}$ Variance of the centralized MVU estimator
- To attain CRLB_{min}, we need $\tau_c = \theta$ (... but θ is unknown!)
- For $\tau_c \in [\theta - \sigma_n, \theta + \sigma_n]$, $\text{Var}\left(\hat{\theta}_{\text{ML}}\right) \leq 2.52\frac{\sigma_n^2}{N}$
- Performance loss wrt centralized & $\tau_c = \theta$

Notes

Problem setup and assumptions:

- Parameter: θ, deterministic and scalar.
- Observations: $x_k = \theta + n_k$ with $k = 1 \dots N$
  - n_k: zero-mean, variance $\sigma_{n,k}^2 = \sigma_k^2$ (different observation qualities!)
  - Noise PDF unknown.
  - Finite dynamic range: $x_k \in [-W, W]$
- Digital sensor-FC communications: uncoded M-QAM modulation
- Uniform L_k-bit quantizers in each sensor node (k)
- BLUE estimator at the FC

Goal: Optimal number of quantization levels $(2^{L_k} - 1)$ in each sensor node?

- Target distortion constraint: $D = \mathbb{E}_x\left[\left(\hat{\theta}(\mathbf{x}) - \theta\right)^2\right] \leq D_0$
- Minimum power consumption.

Notes

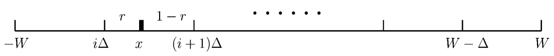
Benchmark: *Centralized* BLUE with *analog* observations

$$\hat{\theta}_{\text{BLUE}} = \left(\sum_{k=1}^{N} \frac{1}{\sigma_k^2}\right)^{-1} \sum_{k=1}^{N} \frac{x_k}{\sigma_k^2}$$

with $\text{Var}(\hat{\theta}) = \left(\sum_{k=1}^{N} \frac{1}{\sigma_k^2}\right)^{-1}$

Quantizer: uniform intervals

- Divide $[-W, W]$ into $2^{L_k} - 1$ intervals, of length $\Delta = 2W/(2^{L_k} - 1)$



- Quantize x_k with an L_k-bit message, $m_k(x_k, L_k)$, while fulfilling $\mathbb{E}\left[m(x_k, L_k)\right] = x_k$

Notes

Case #1: Distortion without errors in sensor-FC channels:

- The message sent is

$$m_k(x_k, L_k) = \theta + n_k + \underbrace{\nu_k}_{\text{Quant. Noise}}$$

with

$$\text{Var}\left(\nu_k\right) \leq \delta_k^2 = \frac{W^2}{\left(2^{L_k} - 1\right)^2}$$

- Quasi-BLUE estimate at the FC:

$$\hat{\theta}_{\text{Q-BLUE}} = \left(\sum_{k=1}^{N} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1} \left(\sum_{k=1}^{N} \frac{m_k}{\sigma_k^2 + \delta_k^2}\right)$$

with distortion lower-bounded by:

$$D = \text{Var}(\hat{\theta}_{\text{Q-BLUE}}) \geq \left(\sum_{k=1}^{N} \frac{1}{\sigma_k^2 + \delta_k^2}\right)^{-1}$$

Notes

Case #2: Distortion with errors in sensor-FC channels:

- Bit Error Rate (BER) of the k-th link: p_b^k
- Received and sent messages differ:

$$m'_k(x_k, L_k) \neq m_k(x_k, L_k) \Longrightarrow \text{increased estimation variance}$$

- Distortion of the quasi-BLUE estimate:

$$D \leq (1 + p_o) \left(\sum_{k=1}^{N} \frac{1}{\sigma_k^2 + \delta_k^2} \right)^{-1}$$

with

$$p_o = \max_{1 \leq k \leq N} \frac{8W}{\sigma_k} \sqrt{\frac{N p_b^k}{3}}$$

Notes

Optimal bit and power allocation:

- For a given distortion target D_o and BER, minimize transmit power
- Optimal number of quantization bits (sensor-wise):

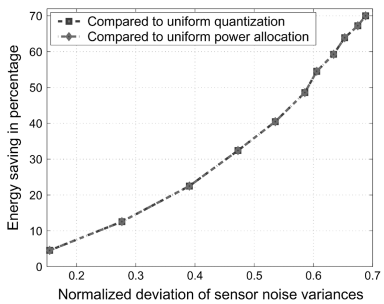
$$L_{\text{opt}}^k = \begin{cases} 0 & \text{if excessive attenuation } a_k \\ \log\left(1 + \frac{W}{\sigma_k}\sqrt{\frac{\eta_0}{a_k}} - 1\right) & \text{otherwise} \end{cases}$$

- Transmit power required:

$$p_k \sim a_k \left(2^{L_k} - 1\right)$$

and $P_{\text{total}} = \sum_{k=1}^{N} p_k$

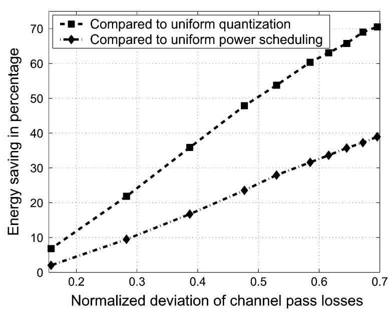
Notes



"Uniform quantization/power allocation": Same number of quantization bits/transmit power in all sensors (to meet D_o)

- The larger the spread in sensor noise variances, the higher the savings.

Notes



"Uniform quantization/power allocation": Same number of quantization bits/transmit power in all sensors (to meet D_o)

- The larger the spread in channel path-losses, the higher the savings.

Notes

Concluding Remarks

- General observations:
  - BLUE and Maximum-Likelihood (ML) estimators are often used in decentralized settings.
  - Benchmarks: centralized estimator and Cramer-Rao Lower Bound.
  - Maximization of sum-rate vs. maximization of estimation accuracy.
- Decentralized Estimation with Analog Transmission (A&F):
  - Opportunistic Power Allocation (OPA) performs close to Optimal Power Allocation (i.e. waterfilling-like solutions) with less complexity and feedback requirements.
  - OPA neatly outperforms Uniform Power Allocation approaches.
- Decentralized Estimation with Digital Transmission:
  - With 1 bit per sensor, ML estimation and rough knowledge on quantization threshold, the distortion is close to the "analog" system.
  - With L_k bits (variable) per sensor, power allocation, and BLUE estimation substantial energy savings w.r.t. the cases with fixed quantization levels/fixed power allocation (for a given distortion).

Notes

References I

Y. Chen and Q. Zhao.
Distributed transmission protocol for lifetime maximization in sensor network.
In *Proc. SPAWC*, pages 895–899, June 2005.

D. Gesbert and M.-S. Alouini.
How much feedback is multi-user diversity really worth?
In *IEEE ICC*, pages 234–238, June 2004.

A. Ribeiro and G. B. Giannakis.
Bandwidth-constrained distributed estimation for wireless sensor networks part i: Gaussian case.
IEEE Trans. on Signal Proc., 54(3):1131–1143, Apr. 2005.

Notes

References II

S. Sanayei and A. Nosratinia.
Opportunistic beamforming with limited feedback.
IEEE Transactions on Wireless Communications, 6(8):2765–2771, 2007.

Notes

THANKS FOR YOUR KIND ATTENTION !!
Questions?

Notes