Channel-Aware Distributed Detection in Wireless Sensor Networks

Bernhard Geiger

25. November 2008

Overview

Introduction

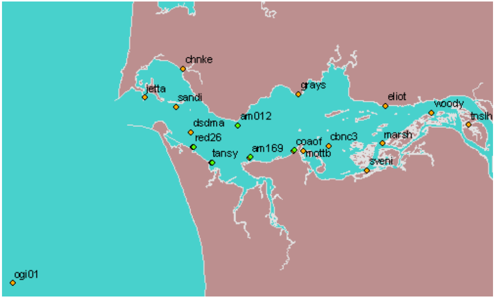
Why is Channel-Awareness necessary?

When is Channel-Awareness necessary?

Channel-Aware Signal Processing
- With complete Channel Knowledge
- With partial Channel Knowledge
- Without Channel Knowledge

Summary

Distributed Detection - CORIE Example



- multiple nodes (K sensors)
- detection of flow/ebb (hypothesis H)
- transmit observations to central node (fusion center)

Classical hypothesis testing with data samples from multiple sensors

Distributed Detection (cont'd)

Two possibilities:

- Recover sensor observations X_k at fusion center: data-centric approach leads to joint source/channel coding problem
- Draw inference on hypothesis H: inference-driven approach leads to channel-aware distributed detection

Distributed Detection (cont'd)

Two possibilities:

- Recover sensor observations X_k at fusion center: data-centric approach leads to joint source/channel coding problem
- Draw inference on hypothesis H: inference-driven approach leads to channel-aware distributed detection

DD usually refers to inference-driven approach (hypothesis testing)!

Distributed Detection (cont'd)

Bandwidth and time constraints: Sensor observations have to be processed in finite time

$$U_k = \gamma_k(X_k)$$

Distributed Detection (cont'd)

Bandwidth and time constraints: Sensor observations have to be processed in finite time

$$U_k = \gamma_k(X_k)$$

Wireless channels suffer from fading, received signals have to be processed as well to decide upon hypothesis

$$U_0 = \gamma_0(\ldots, Y_k, \ldots)$$

Distributed Detection (cont'd)

Two different problems to be solved:

- Design of a fusion rule γ_0 (centralized detection)
- Design of local sensor rules γ_k (source coding, data compression)

For both problems, likelihood ratio (LR) methods are proven to be optimal (even for multilevel quantization).

Distributed Detection (cont'd)

Two different problems to be solved:

- Design of a fusion rule γ_0 (centralized detection)
- Design of local sensor rules γ_k (source coding, data compression)

For both problems, likelihood ratio (LR) methods are proven to be optimal (even for multilevel quantization).

But...

- Local sensor rules (i.e. likelihood ratio thresholds) are coupled with fusion rule, *thus*
- Local sensor rules are coupled with each other

Distributed Detection (cont'd)

Find thresholds via person-by-person optimization (PBPO):

1. Initialize sensor thresholds τ_k
2. $\forall i \in \{0, .., K\}$: For fixed thresholds τ_j $j \in \{0, .., K\}/i$ find optimal τ_i
3. Repeat last step until convergence

Distributed Detection (cont'd)

Find thresholds via person-by-person optimization (PBPO):

1. Initialize sensor thresholds τ_k
2. $\forall i \in \{0, .., K\}$: For fixed thresholds τ_j $j \in \{0, .., K\}/i$ find optimal τ_i
3. Repeat last step until convergence

Effort prohibitively large for WSN with many sensors (large K).
Therefore: Limit considerations to $K = 3$

Overview

Why is Channel-Awareness necessary?

Two levels of uncertainty:

- Source uncertainty (e.g. sensor noise)
- Channel uncertainty (e.g. channel noise, cross-over probability)

Why is Channel-Awareness necessary?

Two levels of uncertainty:

- Source uncertainty (e.g. sensor noise)
- Channel uncertainty (e.g. channel noise, cross-over probability)

Classical DD

- assumes ideal channel (wired),
- computes thresholds considering source noise only *thus*
- separately designs communication schemes → Separation approach

Why is Channel-Awareness necessary?

Two levels of uncertainty:

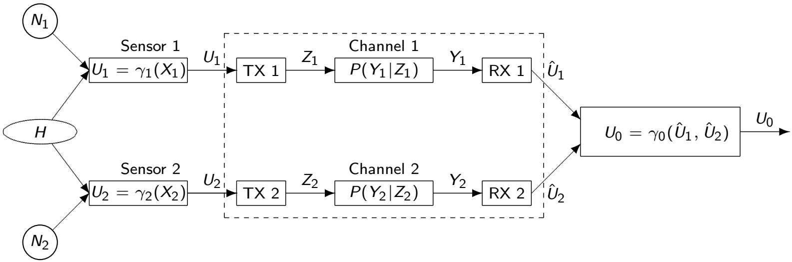
- Source uncertainty (e.g. sensor noise)
- Channel uncertainty (e.g. channel noise, cross-over probability)

Classical DD

- assumes ideal channel (wired),
- computes thresholds considering source noise only *thus*
- separately designs communication schemes → Separation approach

Not sufficient for non-ideal channels!

Example 1



Binary hypothesis, two sensors, iid sensor noise
The communication block is in the dotted box.

Intuitive Explanation

Separation approach:

- Each receiver estimates U_k from Y_k
- Fusion center processes estimates $\hat{U}_k$

Intuitive Explanation

Separation approach:

- Each receiver estimates U_k from Y_k
- Fusion center processes estimates $\hat{U}_k$

Inference-driven approach:

- Better recover H than $U_1 \dots U_K$
- Data Processing Inequality states that fusion center should process Y_k instead
- Otherwise information loss due to quantization $\hat{U}_k = f(Y_k)$

Mathematical Explanation

Assume Example 1 (10) with following parameters:

- Binary hypothesis, i.e.

$$\begin{array}{lll} H_0 & X_k = N_k & \pi_0 = P(H_0) = 0.8 \\ H_1 & X_k = 1 + N_k & \pi_1 = P(H_1) = 0.2 \end{array}$$

- iid sensor noise N_k, Gaussian distributed with $\mu = 0$ and $\sigma^2 = 1$
- Binary symmetric channel (BSC)
- Non-ideal channel with cross-over probabilities $\alpha_1 = 0.05$ and $\alpha_2 = 0.15$

Mathematical Explanation (cont'd)

Let τ_k denote the threshold on X_k observed by sensor k:

Ideal transmission assumption yields

- $\tau_1 = \tau_2 = 0.8474$
- Error probability $P_e = 0.1928$

Mathematical Explanation (cont'd)

Let τ_k denote the threshold on X_k observed by sensor k:

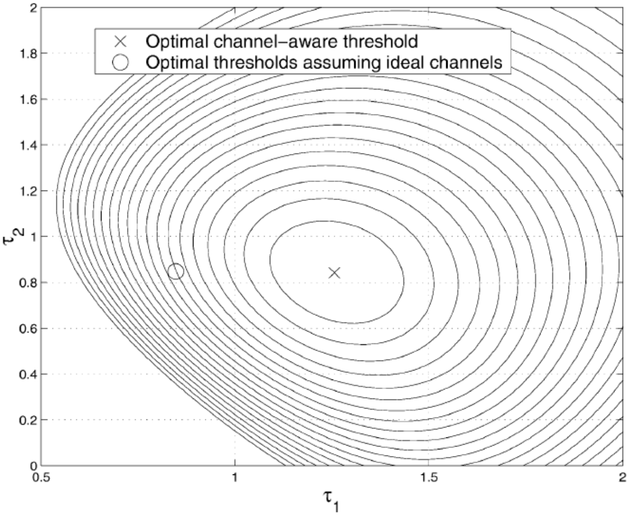
Ideal transmission assumption yields

- $\tau_1 = \tau_2 = 0.8474$
- Error probability $P_e = 0.1928$

Channel-awareness yields

- $\tau_1 = 0.8426$ and $\tau_2 = 1.2570$
- Error probability $P_e = 0.1889$

Mathematical Explanation (cont'd)



Why is Channel-Awareness necessary?

- Channel-aware design improves performance in terms of P_e

Why is Channel-Awareness necessary?

- Channel-aware design improves performance in terms of P_e
- *Or* allows for a higher energy efficiency

Why is Channel-Awareness necessary?

- Channel-aware design improves performance in terms of P_e
- *Or* allows for a higher energy efficiency
- Penalty for assuming ideal channel increases with

Why is Channel-Awareness necessary?

- Channel-aware design improves performance in terms of P_e
- *Or* allows for a higher energy efficiency
- Penalty for assuming ideal channel increases with
  - number of sensors K

Why is Channel-Awareness necessary?

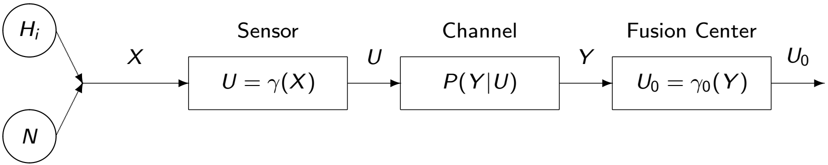
- Channel-aware design improves performance in terms of P_e
- *Or* allows for a higher energy efficiency
- Penalty for assuming ideal channel increases with
  - number of sensors K
  - number of quantization levels M

Overview

When is Channel-Awareness necessary?

Better ask, when it is not...

- Single sensor ($K = 1$)
- Binary hypothesis
- Gaussian noise N



Example 2

Sensor transmits only one bit ($m = 1$) over BSC to fusion center in order to decide on hypothesis.

- Minimize overall error probability
- Take cross-over probability into account

Example 2

Sensor transmits only one bit ($m = 1$) over BSC to fusion center in order to decide on hypothesis.

- Minimize overall error probability
- Take cross-over probability into account

Separate source/channel coding

Proof

Minimum error probability:

$$\begin{aligned} P_e &= \pi_0 P(U_0 = 1|H_0) + \pi_1 P(U_0 = 0|H_1) \\ &= \int_X [P(U = 0|X) D_0(X) + P(U = 1|X) D_1(X)]\, dX \end{aligned}$$

with

$$\begin{aligned} D_0 &= \pi_0 \alpha P(X|H_0) + \pi_1 (1-\alpha) P(X|H_1) \\ D_1 &= \pi_0 (1-\alpha) P(X|H_0) + \pi_1 \alpha P(X|H_1) \end{aligned}$$

Proof (cont'd)

- Choosing threshold τ is like setting $P(U=0|X)=1$ for a set of X (and setting $P(U=1|X)=1$ for the complementary set)

Proof (cont'd)

- Choosing threshold τ is like setting $P(U = 0|X) = 1$ for a set of X (and setting $P(U = 1|X) = 1$ for the complementary set)
- Choose τ so that $D_0(X)$ is min for same set of X (and, at the same time, $D_1(X)$ is min for the complementary set)

Proof (cont'd)

- Choosing threshold τ is like setting $P(U = 0|X) = 1$ for a set of X (and setting $P(U = 1|X) = 1$ for the complementary set)
- Choose τ so that $D_0(X)$ is min for same set of X (and, at the same time, $D_1(X)$ is min for the complementary set)
- Rearrange terms in given equations:

Proof (cont'd)

- Choosing threshold τ is like setting $P(U = 0|X) = 1$ for a set of X (and setting $P(U = 1|X) = 1$ for the complementary set)
- Choose τ so that $D_0(X)$ is min for same set of X (and, at the same time, $D_1(X)$ is min for the complementary set)
- Rearrange terms in given equations:

Choose $i = 0$ for all X where $D_0(X) < D_1(X)$, or:

$$((1-\alpha)-\alpha)\left(\frac{\pi_1 P(X|H_1)}{\pi_0 P(X|H_0)} - 1\right) < 0$$

Proof (cont'd)

For cross-over probabilities less than 50 % the first term falls away.

Proof (cont'd)

For cross-over probabilities less than 50 % the first term falls away.

Only

$$\left(\frac{\pi_1 P(X|H_1)}{\pi_0 P(X|H_0)}\right) < 1$$

remains.

Proof (cont'd)

For cross-over probabilities less than 50 % the first term falls away.

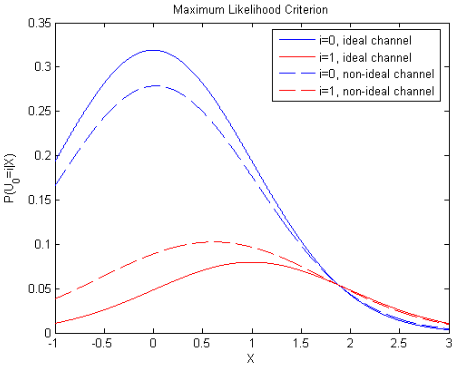
Only

$$\left(\frac{\pi_1 P(X|H_1)}{\pi_0 P(X|H_0)}\right) < 1$$

remains.

Well-known result: MAP

Proof (cont'd)



Simulated threshold for sensor 2 (same parameters as in example 1 (10), i.e. $\alpha_2 = 0.15$)

When is Channel-Awareness necessary?

- *Either* more than one sensor ($K > 1$)
- *Or* M-ary hypothesis ($M > 2$)
- *Or both*

Overview

Channel-Aware Signal Processing

Incorporation of channel state information (CSI) into signal processing at

- Fusion center: centralized detection problem, simple if local sensor rules are given
- Distributed local sensors: strongly affected by CSI (e.g. no channel phase information leads to incoherent transmission schemes and different sensor rules)

Channel-Aware Signal Processing

Incorporation of channel state information (CSI) into signal processing at

- Fusion center: centralized detection problem, simple if local sensor rules are given
- Distributed local sensors: strongly affected by CSI (e.g. no channel phase information leads to incoherent transmission schemes and different sensor rules)

Considering different levels of channel knowledge:

- Complete channel knowledge
- Partial channel knowledge
- No channel knowledge

Overview

Complete Channel Knowledge

Global CSI is available, i.e. we know $\mathbf{h}$. Or, equivalently,

Complete Channel Knowledge

Global CSI is available, i.e. we know **h**. Or, equivalently,

we know the cross-over probabilities for each BSC between sensor k and fusion center, i.e.

$$P(Y_k|U_k)\ \forall k$$

Complete Channel Knowledge

Global CSI is available, i.e. we know $\mathbf{h}$. Or, equivalently,

we know the cross-over probabilities for each BSC between sensor k and fusion center, i.e.

$$P(Y_k|U_k)\ \forall k$$

Only theoretically significant: Lower bounds on P_e

Minimizing Error Probability

Decision rules $\gamma_k(X_k)$ are designed so that

$$P_e(\gamma_0, ..., \gamma_K; \mathbf{h})$$

is minimized.

Minimizing Error Probability

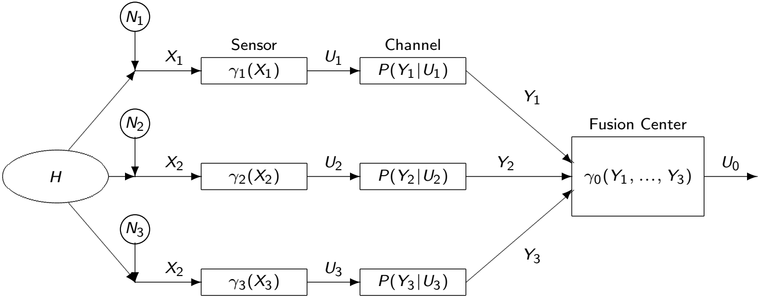
Decision rules $\gamma_k(X_k)$ are designed so that

$$P_e(\gamma_0, ..., \gamma_K; \mathbf{h})$$

is minimized.

Inherent adaptivity: thresholds depend on $\mathbf{h}(t)$, therefore adaptive, cf. Example 1 (10)

Example 3



N_1
N_2
N_3
H
X_1
X_2
X_2
Sensor
$\gamma_1(X_1)$
$\gamma_2(X_2)$
$\gamma_3(X_3)$
U_1
U_2
U_3
Channel
$P(Y_1|U_1)$
$P(Y_2|U_2)$
$P(Y_3|U_3)$
Y_1
Y_2
Y_3
Fusion Center
$\gamma_0(Y_1, \ldots, Y_3)$
U_0

Example 3 (cont'd)

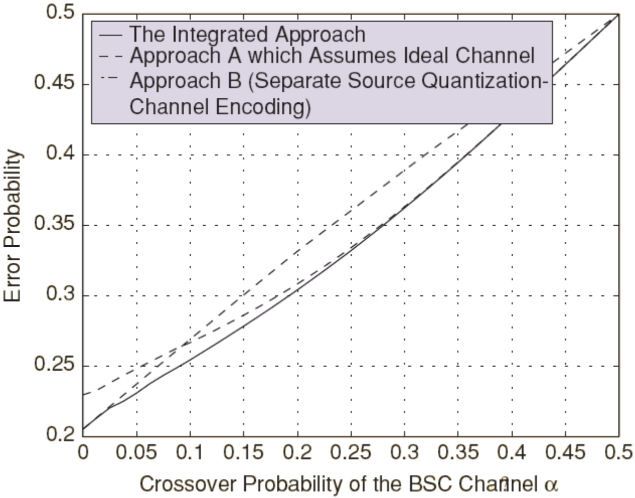
- $K = 3$ sensors
- $m = 2$ bits each sensor (binary hypothesis, soft decision, 3 thresholds)
- α is identical for all channels
- $\pi_0 = \pi_1 = 0.5$

Example 3 (cont'd)

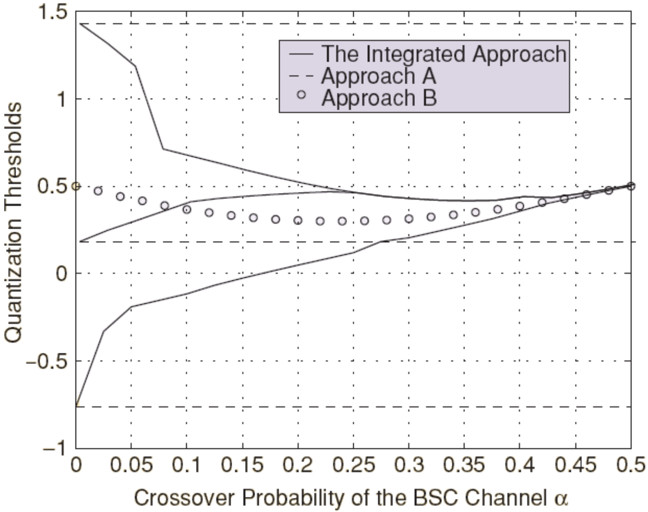
Two approaches are compared to channel-aware design, namely

- Approach A: Ideal channel assumption
- Approach B: Non-ideal channel assumption but separate source/channel coding.
  Here the observation is quantized with $n = 1$ bit, but transmitted using a $m = 2$ bit block code (e.g. repetition code).

Results


The Integrated Approach
Approach A which Assumes Ideal Channel
Approach B (Separate Source Quantization-Channel Encoding)
Error Probability
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
Crossover Probability of the BSC Channel α

Results (cont'd)



Results (cont'd)

- All sensors have same thresholds
- Thresholds for Approach A are constant (channel blind)
- Single Threshold for Approach B
- For a noisy channel, channel-aware approach degrades to Approach

Overview

Partial Channel Knowledge

Global CSI is hardly ever available

- Mobile sensors *and*
- mobile objects in the environment

lead to fading channels.

Partial Channel Knowledge

Global CSI is hardly ever available

- Mobile sensors *and*
- mobile objects in the environment

lead to fading channels.

Reasonable assumption:
Fading statistics are

- stationary *and*
- available

Minimizing Average Error Probability

Decision rules $\gamma_k(X_k)$ should be designed so that

$$\int_h P_e(\gamma_0, ..., \gamma_K; \mathbf{h}) p(\mathbf{h}) d\mathbf{h}$$

is minimized.

Minimizing Average Error Probability

Decision rules $\gamma_k(X_k)$ should be designed so that

$$\int_h P_e(\gamma_0, ..., \gamma_K; \mathbf{h}) p(\mathbf{h}) d\mathbf{h}$$

is minimized.

But...

- P_e is highly non-linear
- Difficult to compute numerically
- Exhaustive search intractable for large K

Minimizing Average Error Probability (cont'd)

Compute average channel first:

$$P(Y_k|U_k) = \int_h P(Y_k|U_k; h_k)p(h_k)dh_k$$

Minimizing Average Error Probability (cont'd)

Compute average channel first:

$$P(Y_k|U_k) = \int_h P(Y_k|U_k; h_k)p(h_k)dh_k$$

Use this averaged transmission channel for the channel-aware design (i.e. assume complete channel knowledge).

- Decrease in performance

Example 4

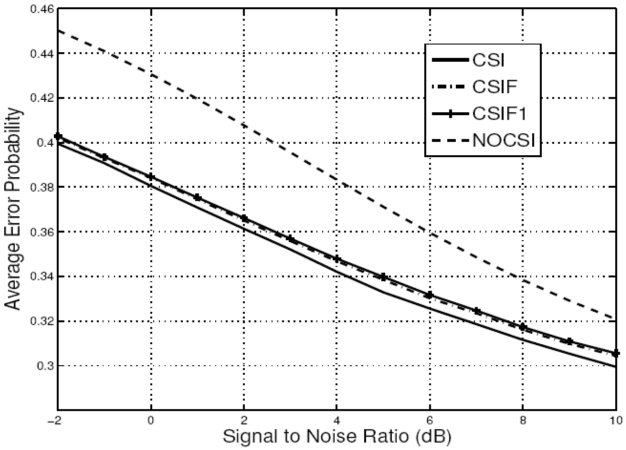
Same setting as in Example 1 (10), but with

- Complex Gaussian sensor noise with $\mu = 0$ and $\sigma^2 = 2$ (ZMCG)
- Rayleigh fading channel

$$Y_k = g_k X_k + W_k$$

with g_k ZMCG with $\sigma_g^2 = 1$ and W_k also ZMCG.

Results


0.46
0.44
0.42
0.4
0.38
0.36
0.34
0.32
0.3
Average Error Probability
–2
0
2
4
6
8
10
Signal to Noise Ratio (dB)
CSI
CSIF
CSIF1
NOCSI

Overview

Introduction

Why is Channel-Awareness necessary?

When is Channel-Awareness necessary?

Channel-Aware Signal Processing

- With complete Channel Knowledge
- With partial Channel Knowledge
- Without Channel Knowledge

Summary

Without Channel Knowledge

- High mobility sensors *and*
- fast moving objects in the environment

lead to fast fading channels. Fading statistics change continuously!

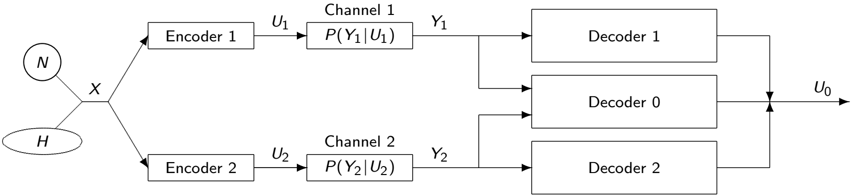
Without Channel Knowledge

- High mobility sensors *and*
- fast moving objects in the environment

lead to fast fading channels. Fading statistics change continuously!

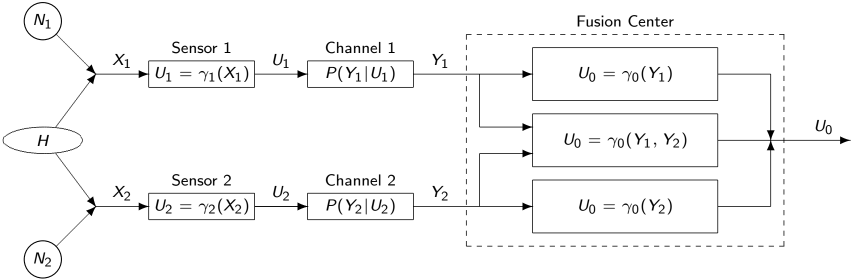
Is channel-aware design possible?

Multiple Description Coding



- Each encoder has same observation
- If one channel information is lost, side decoders ensure acceptable performance

Multiple Description Coding (cont'd)



- Each encoder has its own observation of the phenomenon
- If one channel information is lost...?

Minimizing Error Probability

Assuming ideal channel design $\gamma_k(X_k)$ so that

$$P_e = \pi_0 P(U_0 = 1|H_0) + \pi_1 P(U_0 = 0|H_1)$$

is minimized

Minimizing Error Probability

Assuming ideal channel design $\gamma_k(X_k)$ so that

$$P_e = \pi_0 P(U_0 = 1|H_0) + \pi_1 P(U_0 = 0|H_1)$$

is minimized

Under the constraint that performance remains acceptable if channel information is lost

$$P_e^k = \pi_0 P(U_k = 1|H_0) + \pi_1 P(U_k = 0|H_1) \leq \eta \ \forall k$$

Example 5

Assume binary hypothesis with equal prior probability ($\pi_0 = \pi_1$).
We employ $K = 2$ sensors which observe a ternary variable X_k

$$\begin{array}{ll} P(X_k = 0|H_0) = 0.95 & P(X_k = 0|H_1) = 0.05 \\ P(X_k = 1|H_0) = 0.05 & P(X_k = 1|H_1) = 0.90 \\ P(X_k = 2|H_0) = 0 & P(X_k = 2|H_1) = 0.05 \end{array}$$

and two possible decision rules:

$$\begin{array}{lll} U_k = 0 \text{ if } X_k = 0 & \text{and} & U_k = 1 \text{ else} \\ U_k = 0 \text{ if } X_k \in \{0, 1\} & \text{and} & U_k = 1 \text{ else} \end{array}$$

Results

First rule can be considered as main information, second rule delivers side information only (improve performance).

Classical DD would use different decision rules on each sensor, minimizing overall $P_e = 0.04875$ assuming that channel is reliable.

If one channel information is lost, error probability increases to $P_e = 0.475$!

Results (cont'd)

Increase robustness by using first decision rule on both sensors.

Error probability is limited to $P_e = 0.05$, even if one channel information is lost.

Achievable by constrained minimization (η), but only for small K.

Overview

Summary

Main problem for DD in WSNs is need for data compression.
Parallels to

- distributed source coding
- distributed joint source/channel coding
- CEO problem
- etc.

Summary

Main problem for DD in WSNs is need for data compression.
Parallels to

- distributed source coding
- distributed joint source/channel coding
- CEO problem
- etc.

DD in WSNs is inference-driven, delay and bandwidth constrained and has to cope with fading channels

Summary (cont'd)

Channel-Aware Distributed Detection with

- complete channel knowledge: channel-aware design
- partial channel knowledge: average channel
- no channel knowledge: MDC to increase robustness

Summary (cont'd)

Channel-Aware Distributed Detection with

- complete channel knowledge: channel-aware design
- partial channel knowledge: average channel
- no channel knowledge: MDC to increase robustness

Generalization to large scale networks possible, but hindered by complexity

- complexity/performance trade-offs
- e.g. use same decision rule at each sensor

Outlook

Many problems to be overcome in large scale networks:

- Sensor observations not independend: spatial sampling theoreme?

Outlook

Many problems to be overcome in large scale networks:

- Sensor observations not independend: spatial sampling theoreme?
- Channels not independend: multiple access channel (MAC)

Outlook

Many problems to be overcome in large scale networks:

- Sensor observations not independend: spatial sampling theoreme?
- Channels not independend: multiple access channel (MAC)
- Transmitting decision rules to sensors

Outlook

Many problems to be overcome in large scale networks:

- Sensor observations not independend: spatial sampling theoreme?
- Channels not independend: multiple access channel (MAC)
- Transmitting decision rules to sensors
- Computational complexity

Outlook

Many problems to be overcome in large scale networks:

- Sensor observations not independend: spatial sampling theoreme?
- Channels not independend: multiple access channel (MAC)
- Transmitting decision rules to sensors
- Computational complexity
- Networking issues...

That's it!

Thanks for your attention!