Channel-aware Distributed Classification in Wireless Sensor Networks Using Binary Local Decisions

Mohammad Fanaei, Matthew C. Valenti, Natalia A. Schmid, and Vinodkrishnan Kulathumani

Wireless Communications Research Lab (WCRL)
Department of Computer Science and Electrical Engineering
West Virginia University, Morgantown, WV, USA

2011 SPIE Defense, Security, and Sensing
Signal Processing, Sensor Fusion, and Target Recognition XX
April 27, 2011

Outline

1. Introduction, Motivation, and Problem Statement
2. System Model of Our Distributed Classification WSN
3. Fusion Rule Derivation
4. Numerical Analysis
5. Conclusions

Outline

1 Introduction, Motivation, and Problem Statement

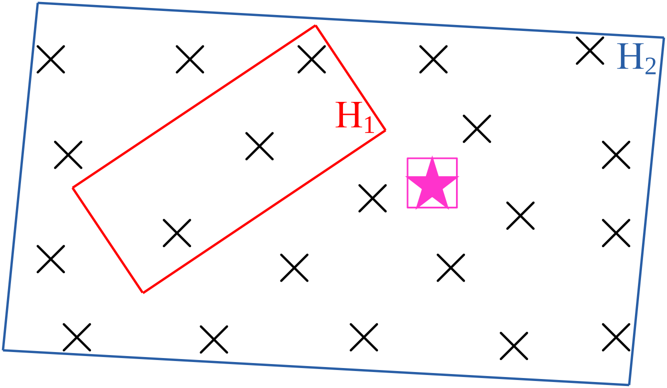
2 System Model of Our Distributed Classification WSN

3 Fusion Rule Derivation

4 Numerical Analysis

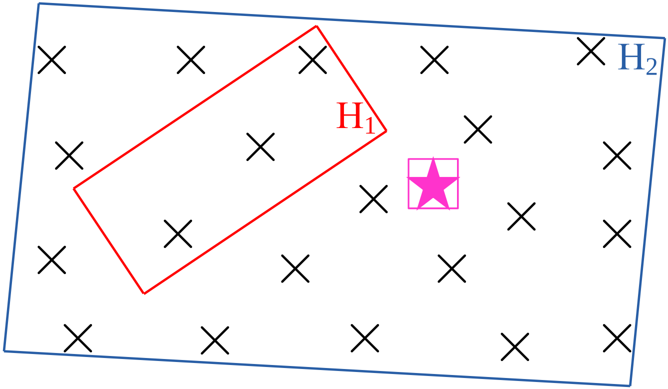
5 Conclusions

WSNs and Distributed Detection and Classification (DDC)



- Wireless sensor network (WSN): Large number of densely-deployed spatially-distributed autonomous sensors with limited capabilities that cooperate with each other to achieve a common goal.

WSNs and Distributed Detection and Classification (DDC)



- Main constraints:
  - Limited bandwidth and energy resources.
  - Limited processing capability of local sensors.

Problem Statement and Possible Solutions

- **General Problem of DDC in WSNs:**
  - Distributed local sensors observe the conditions of their surrounding environment and process their local observations.
  - The processed data is sent to a fusion center (FC).
  - The FC makes the ultimate global decision.

- **Possible Approaches:** There are M hypotheses to be classified.
  - Multi-class classification by each sensor and sending it with $\lceil \log_2 M \rceil$ bits.
  - Multi-class classification by each sensor and sending it with less bits than $\lceil \log_2 M \rceil$ (for example, binary decision transmission).
  - Binary classification by each sensor and sending binary decisions.

Problem Statement and Possible Solutions

- **General Problem of DDC in WSNs:**
  - Distributed local sensors observe the conditions of their surrounding environment and process their local observations.
  - The processed data is sent to a fusion center (FC).
  - The FC makes the ultimate global decision.

- **Possible Approaches:** There are M hypotheses to be classified.
  - Multi-class classification by each sensor and sending it with $\lceil \log_2 M \rceil$ bits.
  - Multi-class classification by each sensor and sending it with less bits than $\lceil \log_2 M \rceil$ (for example, binary decision transmission).
  - Binary classification by each sensor and sending binary decisions.

Outline

1 Introduction, Motivation, and Problem Statement

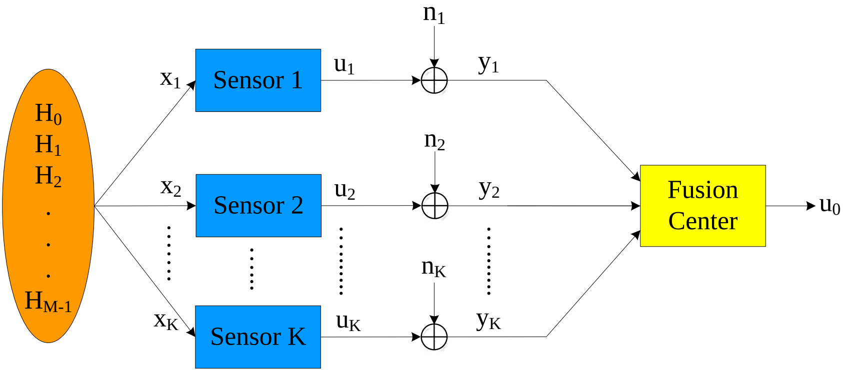
2 System Model of Our Distributed Classification WSN

3 Fusion Rule Derivation

4 Numerical Analysis

5 Conclusions

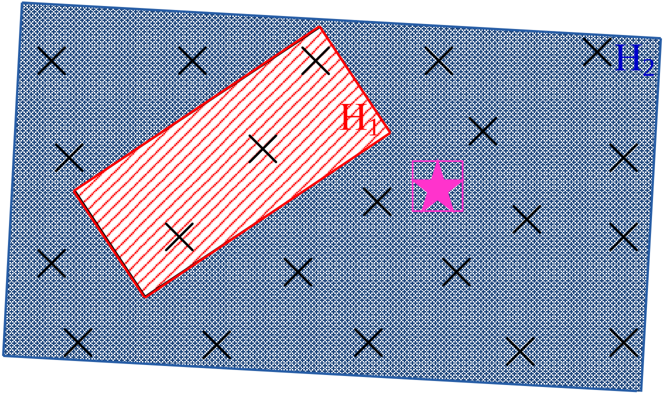
System Model of Our Distributed Classification WSN I



- M: The number of independent and mutually exclusive hypotheses, $M \geq 2$, with *known* prior probabilities of $p_j = P[H = H_j]$.
  - H_0 is the null/rejection hypothesis and its existence means that none of the other $M-1$ hypotheses has occurred.

System Model of Our Distributed Classification WSN II

- Each non-null hypothesis H_j is associated with a known *influence field*, A_j.



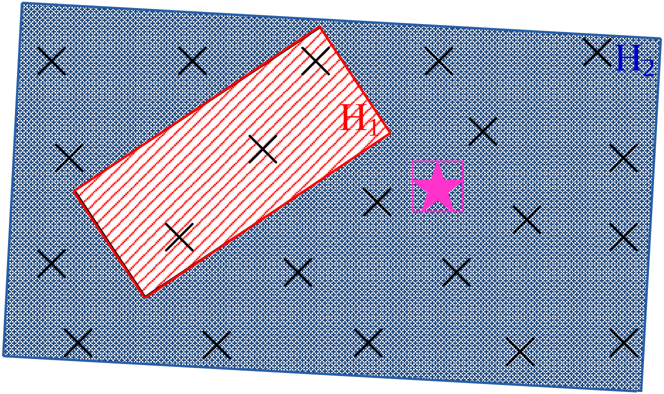
System Model of Our Distributed Classification WSN II

- Each non-null hypothesis H_j is associated with a known *influence field*, A_j.
  - Spatial region in the surrounding of H_j in which it can be sensed using some specific modality.
  - Each influence field is characterized by its area and considered flat over its area.
- Entire influence field of the underlying hypothesis is inside the sensing area.
- The center of the influence field of the underlying hypothesis is known or has been reliably estimated

System Model of Our Distributed Classification WSN II

- If the sensors are distributed uniformly within the sensing area, the average number of sensors that can be placed in the influence field of hypothesis H_j will be $K_j = \lfloor \frac{A_j}{S} K \rfloor$.

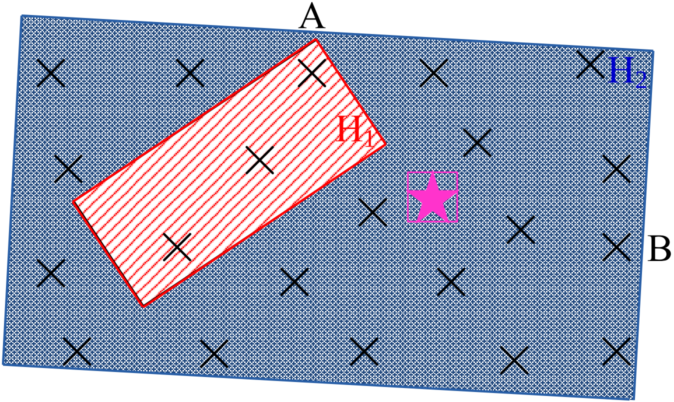
$$K_0 = 0 \qquad K_1 = 3 \qquad K_2 = 21$$



System Model of Our Distributed Classification WSN II

- Assuming *known* center of underlying influence field, for each sensor i the set of hypotheses is divided into two disjoint subsets: $\mathcal{C}_i^1$ and $\mathcal{C}_i^0$.

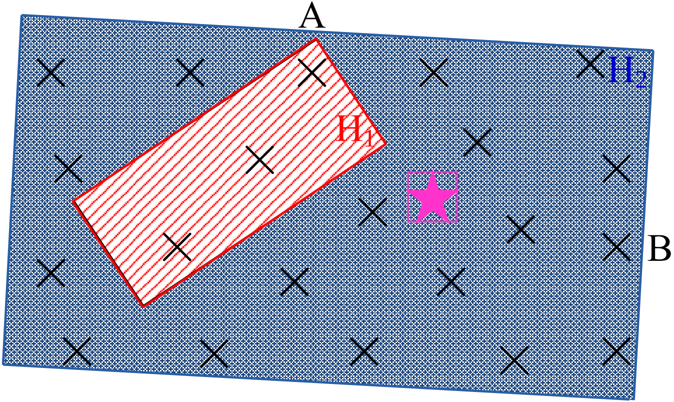
$$\mathcal{C}_A^1 = \{H_1, H_2\} \qquad \mathcal{C}_A^0 = \{H_0\}$$



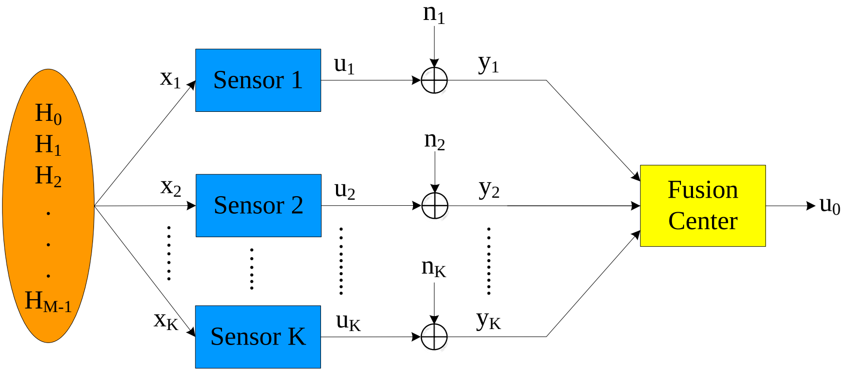
System Model of Our Distributed Classification WSN II

- Assuming *known* center of underlying influence field, for each sensor i the set of hypotheses is divided into two disjoint subsets: $\mathcal{C}_i^1$ and $\mathcal{C}_i^0$.

$$\mathcal{C}_B^1 = \{H_2\} \qquad \mathcal{C}_B^0 = \{H_0, H_1\}$$

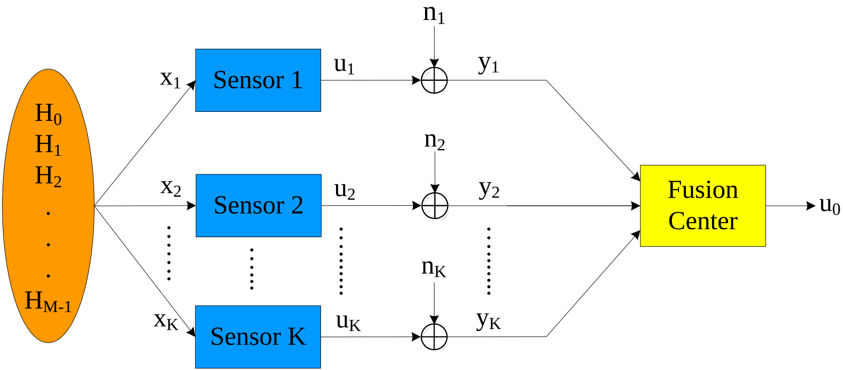


System Model of Our Distributed Classification WSN III



- K: The number of distributed local sensors.
- $\mathbf{x} = [X_1, X_2, \ldots, X_K]$: Vector of independent (given any specific underlying hypothesis) local sensor observations: $p(\mathbf{x}|H_j) = \prod_{i=1}^{K} p(X_i|H_j)$.

System Model of Our Distributed Classification WSN III

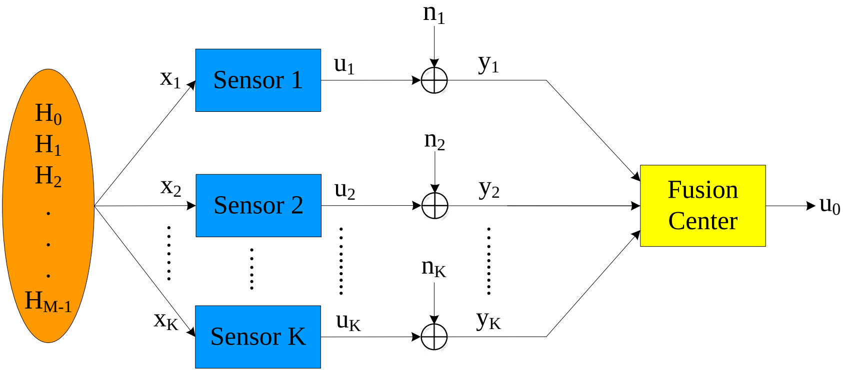


- Conditional local observation of sensor i, given hypothesis H_j:

$$X_i|H_j = \begin{cases} v_i, & H_j \in \mathcal{C}_i^0 \\ s + v_i, & H_j \in \mathcal{C}_i^1 \end{cases}$$

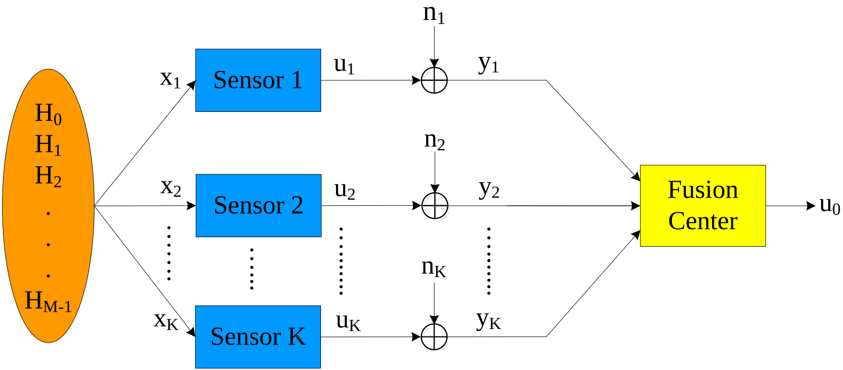
with i.i.d. $v_i \sim \mathcal{N}(0, \sigma_o^2)$.

System Model of Our Distributed Classification WSN III



- U_i, $1 \leq i \leq K$: Binary local decision based on following rule:
$$U_i = \gamma_i(X_i) = \begin{cases} 0, & X_i < \beta_i \\ 1, & X_i > \beta_i \end{cases}$$
- Sensors are able to distinguish only the occurrence or not occurrence of the $M-1$ non-null hypotheses.

System Model of Our Distributed Classification WSN III



- U_i's are sent to the fusion center through parallel transmission channels.
- $\mathbf{y} = [Y_1, Y_2, \ldots, Y_K]$: The output of the parallel AWGN channel received by the fusion center.
  - $Y_i = U_i + n_i$ where i.i.d. $n_i \sim \mathcal{N}(0, \sigma_n^2)$.
- $U_0 = \gamma_0(\mathbf{y}) \in \{0, 1, \ldots, M-1\}$: The final decision at the fusion center.

Outline

1 Introduction, Motivation, and Problem Statement

2 System Model of Our Distributed Classification WSN

3 Fusion Rule Derivation

4 Numerical Analysis

5 Conclusions

Fusion Center's Decision Metric

- $S_0 = \sum_{i=1}^{K} Y_i$: Fusion center's *decision metric*.

$$S_0 = \sum_{i=1}^{K} U_i + n_0, \quad n_0 \sim \mathcal{N}(0, K\sigma_n^2)$$

  - S_0 tends to have larger values if the influence field of the underlying hypothesis is larger.
  - Influence fields of different hypotheses should have enough separation.

- $\Gamma = \{\alpha_1, \alpha_2, \ldots, \alpha_{M-1}\}$: The set of decision thresholds based on which the fusion center can classify the underlying hypothesis using S_0. The fusion rule will be as follows: $(\alpha_0 = -\infty,\ \alpha_M = \infty)$

$$U_0 = j \ \text{ if and only if } \ \alpha_j \leq S_0 < \alpha_{j+1}, \quad j = 0, 1, \ldots, M-1$$

Overview of Fusion Rule Derivation I

Find conditional pdf of sensor decision, U_i, given hypothesis H_j.
$f_{U_i|H_j}(u_i|H_j)$

←

$$X_i|H_j = \begin{cases} v_i, & \text{if } H_j \in \mathcal{C}_i^0 \\ s + v_i, & \text{if } H_j \in \mathcal{C}_i^1 \end{cases}$$

$v_i \sim \mathcal{N}(0, \sigma_o^2)$ AND

$$u_i = \gamma_i(x_i) = \begin{cases} 0, & \text{if } x_i < \beta_i \\ 1, & \text{if } x_i > \beta_i \end{cases}$$

Overview of Fusion Rule Derivation I

Find conditional pdf of sensor decision, U_i, given hypothesis H_j.
$f_{U_i|H_j}(u_i|H_j)$

$$X_i|H_j = \begin{cases} v_i, & \text{if } H_j \in \mathcal{C}_i^0 \\ s + v_i, & \text{if } H_j \in \mathcal{C}_i^1 \end{cases}$$

$v_i \sim \mathcal{N}(0, \sigma_o^2)$ AND

$$u_i = \gamma_i(x_i) = \begin{cases} 0, & \text{if } x_i < \beta_i \\ 1, & \text{if } x_i > \beta_i \end{cases}$$

Find conditional pdf of received signal at fusion center, Y_i, given decision at local sensor, U_i.
$f_{Y_i|U_i}(y_i|u_i)$

$y_i = u_i + n_i$, $n_i \sim \mathcal{N}(0, \sigma_n^2)$

Overview of Fusion Rule Derivation I

$$X_i|H_j = \begin{cases} v_i, & \text{if } H_j \in \mathcal{C}_i^0 \\ s + v_i, & \text{if } H_j \in \mathcal{C}_i^1 \end{cases}$$

$$v_i \sim \mathcal{N}(0, \sigma_o^2) \text{ AND}$$

$$u_i = \gamma_i(x_i) = \begin{cases} 0, & \text{if } x_i < \beta_i \\ 1, & \text{if } x_i > \beta_i \end{cases}$$

Find conditional pdf of sensor decision, U_i, given hypothesis H_j.
$f_{U_i|H_j}(u_i|H_j)$

$$y_i = u_i + n_i,\ n_i \sim \mathcal{N}(0, \sigma_n^2)$$

Find conditional pdf of received signal at fusion center, Y_i, given decision at local sensor, U_i.
$f_{Y_i|U_i}(y_i|u_i)$

$$f_{Y_i|H_j} = f_{Y_i|U_i} \times f_{U_i|H_j}$$

Overview of Fusion Rule Derivation I

Find conditional pdf of sensor decision, U_i, given hypothesis H_j.
$f_{U_i|H_j}(u_i|H_j)$

Find conditional pdf of received signal at fusion center, Y_i, given decision at local sensor, U_i.
$f_{Y_i|U_i}(y_i|u_i)$

$$f_{Y_i|H_j} = f_{Y_i|U_i} \times f_{U_i|H_j}$$

$\mathcal{L}|_{-\nu}$

Find conditional moment generating function (MGF) of Y_i, given H_j.

$$\begin{aligned}\Phi_{Y_i|H_j}(\nu) &= \mathbb{E}\left[e^{\nu Y_i}\middle|H_j\right] \\ &= \int_{-\infty}^{\infty} e^{\nu y_i} f_{Y_i|H_j}(y_i)\, dy_i \\ &= \mathcal{L}\left\{f_{Y_i|H_j}(y_i|H_j)\right\}\Big|_{-\nu}\end{aligned}$$

Definition

Overview of Fusion Rule Derivation I

Find conditional pdf of sensor decision, U_i, given hypothesis H_j.
$f_{U_i|H_j}(u_i|H_j)$

↓

Find conditional pdf of received signal at fusion center, Y_i, given decision at local sensor, U_i.
$f_{Y_i|U_i}(y_i|u_i)$

↓

$$f_{Y_i|H_j} = f_{Y_i|U_i} \times f_{U_i|H_j}$$

$\xrightarrow{\mathcal{L}|_{-\nu}}$

Find conditional moment generating function (MGF) of Y_i, given H_j.

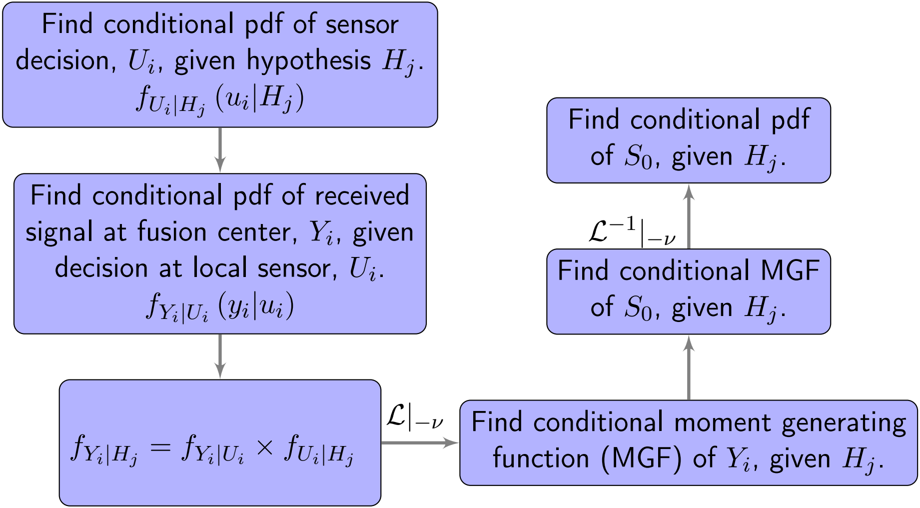
↑

Find conditional MGF of S_0, given H_j.

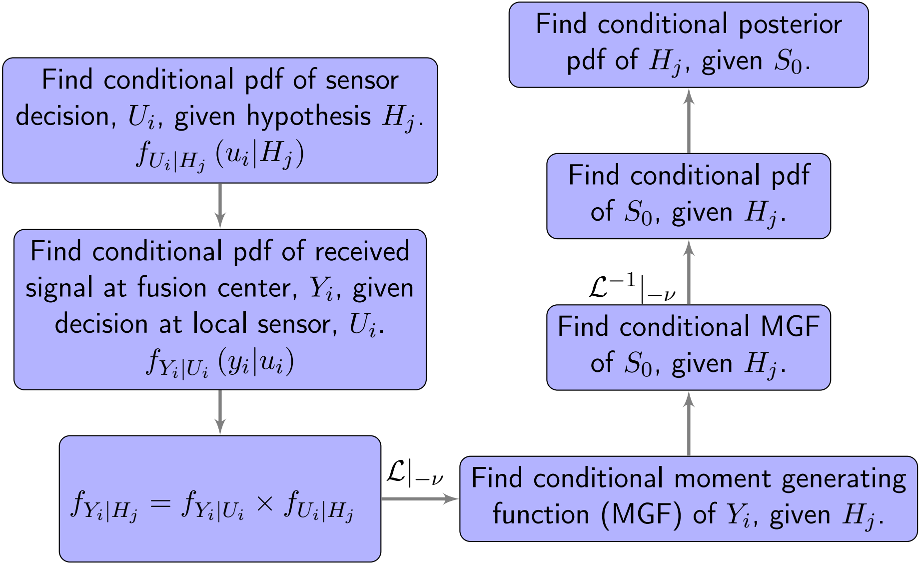
↑

$$\Phi_{S_0|H_j}(\nu) = \prod_{i=1}^{K} \Phi_{Y_i|H_j}(\nu)$$

Overview of Fusion Rule Derivation I



Overview of Fusion Rule Derivation I



Overview of Fusion Rule Derivation II

- Conditional posterior pdf of H_j, given S_0:

$$f_{H_j|S_0}(H_j|s_0) = \frac{p_j}{\sqrt{2\pi K\sigma_n^2}} \sum_{\ell=0}^{K} a_\ell \exp\left[-\frac{(s_0-\ell)^2}{2K\sigma_n^2}\right]$$

- Each coefficient a_ℓ is a specific function of a subset of β_i's, i.e. local sensor decision thresholds.

- Minimum error probability Bayesian decision rule:

$$\hat{H}_j = \underset{j\in\{0,1,\ldots,M-1\}}{\text{arg max}} \; f_{H_j|S_0}(H_j|s_0)$$

- The intersection of $f_{H_j|S_0}(H_j|x)$, $j = 0, 1, \ldots, M-1$, determines the optimum decision region thresholds at the fusion center.

Overview of Fusion Rule Derivation II

- Conditional posterior pdf of H_j, given S_0:

$$f_{H_j|S_0}\left(H_j|s_0\right) \quad = \quad \frac{p_j}{\sqrt{2\pi K\sigma_n^2}} \sum_{\ell=0}^{K} a_\ell \exp\left[-\frac{\left(s_0-\ell\right)^2}{2K\sigma_n^2}\right]$$

- Each coefficient a_ℓ is a specific function of a subset of β_i's, i.e. local sensor decision thresholds.

- Minimum error probability Bayesian decision rule:

$$\hat{H}_j \quad = \quad \underset{j\in\{0,1,\ldots,M-1\}}{\text{arg max}} \; f_{H_j|S_0}\left(H_j|s_0\right)$$

- The intersection of $f_{H_j|S_0}\left(H_j|x\right)$, $j=0,1,\ldots,M-1$, determines the optimum decision region thresholds at the fusion center.

Locally Optimal vs. Globally Optimal Local Thresholds

- Based on Bayesian decision theory, *locally* optimal decision rule of sensor i for a *binary* decision making is

$$P\left[H_j \in \mathcal{C}_i^1 | x_i\right] \underset{u_i=0}{\overset{u_i=1}{\gtrless}} P\left[H_j \in \mathcal{C}_i^0 | x_i\right]$$

$$x_i \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \beta_{i,\text{Local}} = \frac{s}{2} + \frac{\sigma_O^2}{s} \ln\left(\frac{P\left[H_j \in \mathcal{C}_i^0\right]}{P\left[H_j \in \mathcal{C}_i^1\right]}\right)$$

- $\beta_{i,\text{Local}}$ depends only on the variance of observation noise, σ_O^2.

- Globally optimal local decision thresholds depend on variances of both observation noise and channel noise, σ_O^2 and σ_n^2. It is optimized through brute force search over all possible discrete values for β_i's.

Locally Optimal vs. Globally Optimal Local Thresholds

- Based on Bayesian decision theory, *locally* optimal decision rule of sensor i for a *binary* decision making is

$$P\left[H_j \in \mathcal{C}_i^1 | x_i\right] \underset{u_i=0}{\overset{u_i=1}{\gtrless}} P\left[H_j \in \mathcal{C}_i^0 | x_i\right]$$

$$x_i \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \beta_{i,\text{Local}} = \frac{s}{2} + \frac{\sigma_O^2}{s} \ln\left(\frac{P\left[H_j \in \mathcal{C}_i^0\right]}{P\left[H_j \in \mathcal{C}_i^1\right]}\right)$$

- $\beta_{i,\text{Local}}$ depends only on the variance of observation noise, σ_O^2.

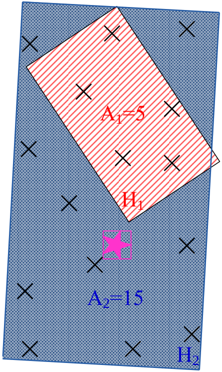
- Globally optimal local decision thresholds depend on variances of both observation noise and channel noise, σ_O^2 and σ_n^2. It is optimized through brute force search over all possible discrete values for β_i's.

Outline

1. Introduction, Motivation, and Problem Statement
2. System Model of Our Distributed Classification WSN
3. Fusion Rule Derivation
4. Numerical Analysis
5. Conclusions

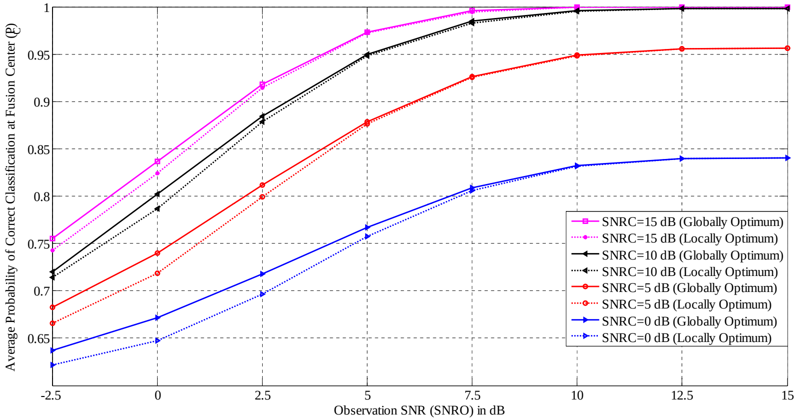
WSN Environment Description

- $K = 15$, $M = 3$ (H_0, H_1, and H_2).
- $P[H_0] = 0.6$, $P[H_1]0.3$, and $P[H_2] = 0.1$.
- Total area of the observation environment = 15.
- A_1 = 5, and A_2 = 15.
- Assuming uniform distribution of the sensors:
  - $K_1 = 5$ and $K_2 = 15$.
- There are two non-null hypotheses.
- Local sensors are divided into two disjoint groups:
  - $K_1 = 5$ sensors that can be inside the influence field of either of the non-null hypotheses.
  - Other $K - K_1 = 10$ sensors that can only be inside the influence field of hypothesis H_2.

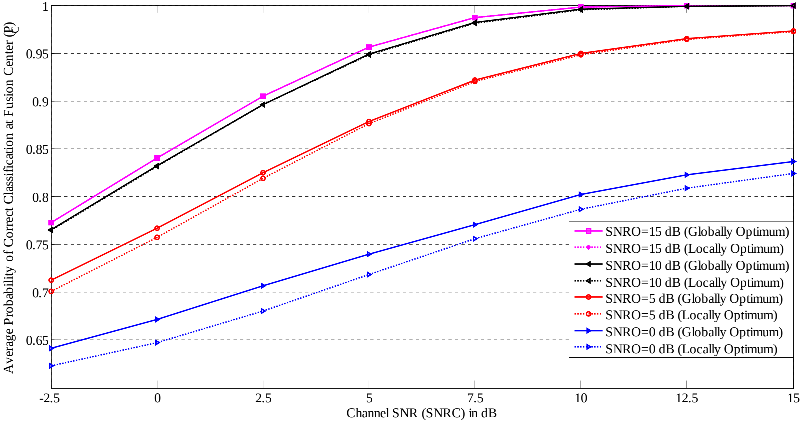


- P_c: Average optimized probability of correct classification at the fusion center.

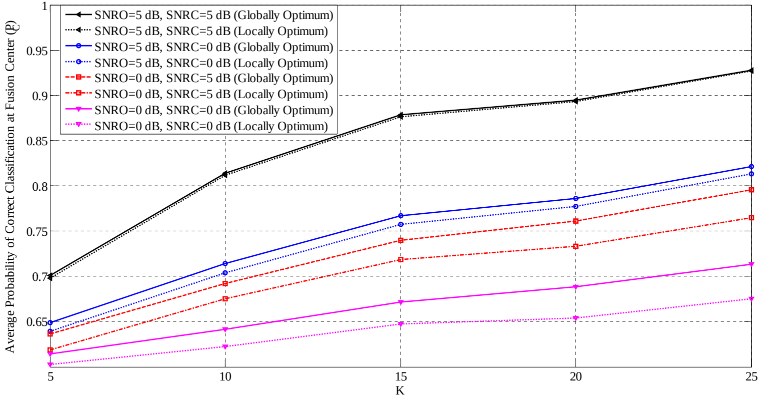
Effect of Observation SNR on Classification Performance



Effect of Channel SNR on Classification Performance



Effect of Number of Sensors on Classification Performance



Outline

1 Introduction, Motivation, and Problem Statement

2 System Model of Our Distributed Classification WSN

3 Fusion Rule Derivation

4 Numerical Analysis

5 Conclusions

Summary and Future Work

Summary

- The problem of multi-class classification using local binary decisions was summarized.
- A model for a simple WSN deployed as a multi-class classification system was introduced.
- Detailed analysis of the proposed classification WSN was presented.
- Numerical performance of the proposed classification system was presented in a typical scenario.

Future Work

- Considering asynchronous communication between the local sensors and fusion center.
- Considering fading channels between the local sensors and fusion center.
- Considering more complex influence fields of the underlying hypothesis.
- Simulation of the proposed distributed classification system using actual data.

Thank You Very Much for Your Attention.

Questions?