Principles of Wireless Sensor Networks

https://www.kth.se/social/course/EL2745/

Lecture 7

Distributed Detection

Carlo Fischione
Associate Professor of Sensor Networks
e-mail:carlofi@kth.se
http://www.ee.kth.se/∼carlofi/

KTH Royal Institute of Technology
Stockholm, Sweden

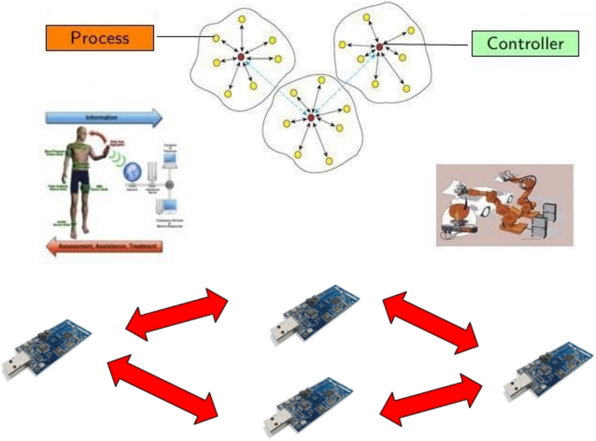
September 23, 2014

Course content

- Part 1
  - Lec 1: Introduction to WSNs
  - Lec 2: Introduction to Programming WSNs
- Part 2
  - Lec 3: Wireless Channel
  - Lec 4: Physical Layer
  - Lec 5: Medium Access Control Layer
  - Lec 6: Routing
- Part 3
  - Lec 7: Distributed Detection
  - Lec 8: Static Distributed Estimation
  - Lec 9: Dynamic Distributed Estimation
  - Lec 10: Positioning and Localization
  - Lec 11: Time Synchronization
- Part 4
  - Lec 12: Wireless Sensor Network Control Systems 1
  - Lec 13: Wireless Sensor Network Control Systems 2
  - Lec 14: Summary and Project Presentations

Previous lecture

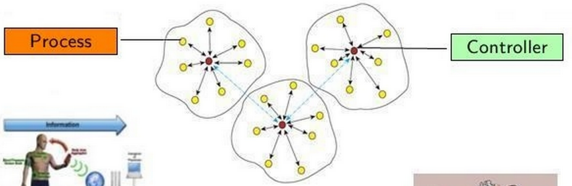
- Application
- Presentation
- Session
- Transport
- Routing
- MAC
- Phy

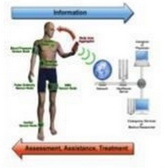


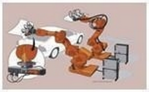
On which path messages should be routed?

Today's lecture

- Application
- Presentation
- Session
- Transport
- Routing
- MAC
- Phy







- Today we study how to detect events out of uncertain (noisy) observations
- Detection is an application on top of the protocol stack
- However, detection theory can be used in other layers as well

Today's learning goals

- What is binary detection?
- How to detect events from one sensor?
- How to detect events from multiple sensors?

Outline

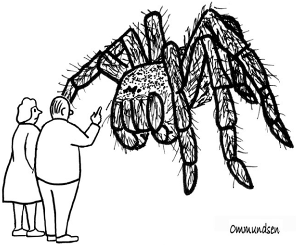
- Introduction to detection theory
- Detection from one sensor
- Detection from multiple sensors

Outline

- Introduction to detection theory
- Detection from one sensor
  - Decision rules, MAP, LRT, ML
  - The Neyman-Pearson criterion
- Detection from multiple sensors
  - The Likelihood Ratio Test
  - The counting rule

Basic of detection theory

Hypothesis concept



"I've narrowed it to two hypotheses: it grew or we shrunk."

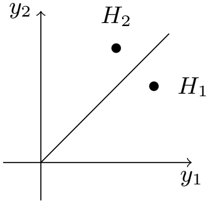
- The concept of binary hypothesis testing is now introduced

H_0	hypothesis 0	s_0 signal
H_1	hypothesis 1	s_1 signal

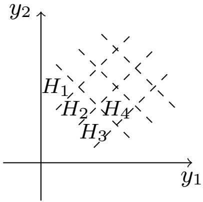
- We consider a measurement $y(t)$: noisy signal associated to the event. Consider $y(t)$ as random variable. Let y be a specific outcome of this random variable

Example: Basic of detection theory

Suppose $y(t) \in \mathbb{R}^2$



- Binary case of hypothesis:
  the signal may fall in two (binary) different areas



- Multiple hypotheses:
  the signal may fall in multiple different areas

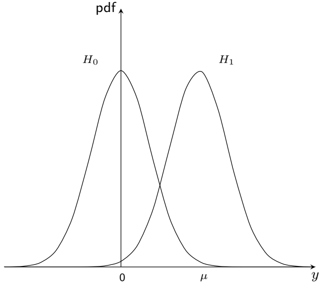
Goal of detection: Minimize the errors out of noisy measurements that may misplace the points

Binary hypothesis testing

- Suppose $y(t) = \begin{cases} s_0 & \text{if } H_0 \text{ happened,} \\ s_1 & \text{if } H_1 \text{ happened.} \end{cases}$

Example: Assume $y(t)$ is simply given by

$$y(t) = \begin{cases} n(t) & \text{if } H_0 \text{ happened,} \\ \mu + n(t) & \text{if } H_1 \text{ happened.} \end{cases}$$



Outline

- Introduction to detection theory
- Detection from one sensor
  - Decision rules, MAP, LRT, ML
  - The Neyman-Pearson criterion
- Detection from multiple sensors
  - The Likelihood Ratio Test
  - The counting rule

Definition of probabilities

$\Pr(s_1|H_0)$ Probability of false alarm

$\Pr(s_0|H_1)$ Probability of miss detection

$\Pr(s_1|H_1)$ Probability of detection

Definition

A posteriori probability: Given the realization of $y(t) : y$, what is the probability that H_0 or H_1 happened?

- $\Pr(H_0|y)$
- $\Pr(H_1|y)$

Criterion of *a posteriori*

- Maximum *a posteriori* probability (MAP)

$$\text{We decide for } \begin{cases} H_0 & \text{if } \Pr(H_0|y) > \Pr(H_1|y) \\ H_1 & \text{if } \Pr(H_0|y) \leq \Pr(H_1|y) \end{cases}$$

- In practice, we assume to know the probabilites $\Pr(H_0)$ and $\Pr(H_1)$. According to Bayes' rule

$$\Pr(H_0|y) = \frac{\Pr(y|H_0)\Pr(H_0)}{\Pr(y)} \qquad \Pr(H_1|y) = \frac{\Pr(y|H_1)\Pr(H_1)}{\Pr(y)}$$

- Therefore MAP criterion becomes

$$\text{We decide for } \begin{cases} H_0 & \text{if } \Pr(y|H_0)\cdot\Pr(H_0) > \Pr(y|H_1)\cdot\Pr(H_1) \\ H_1 & \text{if } \Pr(y|H_0)\cdot\Pr(H_0) \leq \Pr(y|H_1)\cdot\Pr(H_1) \end{cases}$$

Likelihood ratio test

- The previous test can be equivalently converted into the Likelihood Ratio Test (LRT)

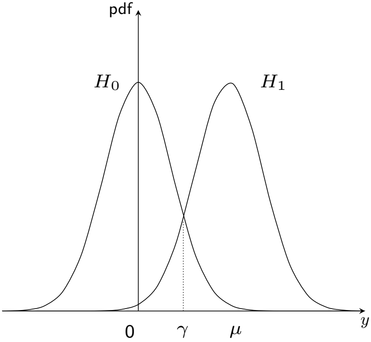
$$\frac{\Pr(y|H_1)}{\Pr(y|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\Pr(H_0)}{\Pr(H_1)}$$

- In the special case where $\dfrac{\Pr(H_0)}{\Pr(H_1)} = 1 \;\Rightarrow$ Maximum Likelihood detection (ML)

Example

Consider again the example on page 9

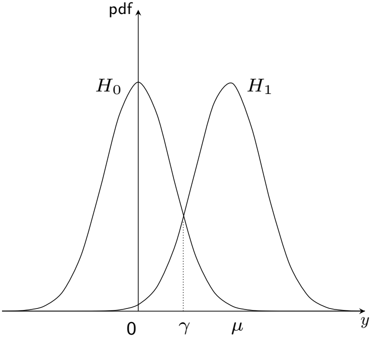
$$y(t) = \begin{cases} n(t) & \text{if } H_0 \text{ happened,} \\ \mu + n(t) & \text{if } H_1 \text{ happened.} \end{cases} \qquad n(t) \in G(0,1)$$



- If $y > \gamma$ we decide for H_1
- If $y \leq \gamma$ we decide for H_0

The intersection point of the two adjacent gaussian curves defines the threshold γ, according to which a decision is made

Example



The conditional probabilities given the hypotheses H_0 and H_1 are respectively

$$\Pr\left(y|H_0\right) = \int_{-\infty}^{\gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \triangleq \Pr\left(H_0|H_0\right)$$

$$\Pr\left(y|H_1\right) = \int_{\gamma}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2}} dz \triangleq \Pr\left(H_1|H_1\right)$$

Probabilites depend on γ

- Probability of false alarm

$$\mathrm{P}_F(\gamma) \triangleq \Pr(H_1|H_0) = \int_{\gamma}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = Q(\gamma)$$

- Probability of miss detection

$$\mathrm{P}_M(\gamma) \triangleq \Pr(H_0|H_1) = \int_{-\infty}^{\gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2}} dz = Q(\mu - \gamma)$$

How should γ be chosen?

Optimization problem

- Tradeoff between γ, P_F and P_M
- Therefore, the choice of γ can be put in form of an optimization problem

There are at least two optimization methods that can be used in order to choose γ

- Fast-Lipschitz optimization
- Pareto optimization

Fast-Lipschitz optimization

Fast-Lipschitz optimization

$$\min_{\gamma} \ \mathrm{P}_M(\gamma) = Q(\mu - \gamma)$$
$$\text{s.t.} \ \ \mathrm{P}_F(\gamma) \leq \bar{\mathrm{P}}_F$$

$\Downarrow$ Solution

$$\gamma^* : \ Q(\gamma^*) = \bar{\mathrm{P}}_F$$

In case that $\Pr(H_0)$, $\Pr(H_1)$ are unknown, we use this method for choosing γ (Neyman-Pearson criterion)

Pareto optimization

Another way to choose the detection threshold, γ, is by using the Pareto optimization method

Pareto optimization

The optimal γ, γ^*, is found by the minimization of the average cost function $C(\gamma)$

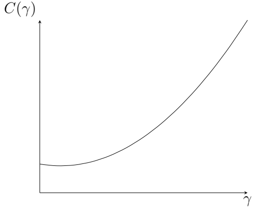
$$\text{Find } \gamma^* \rightarrow \min_{\gamma} C(\gamma)$$

and

$$C(\gamma) = \sum_{i=0}^{1} \sum_{\substack{j=0 \\ j \neq i}}^{1} c_{ij} \Pr(H_i | H_j) \Pr(H_j)$$

where $\Pr(H_i | H_j)$ is either the probability of false alarm or miss detection and c_{ij} the cost of being wrong

Example of $C(\gamma)$:

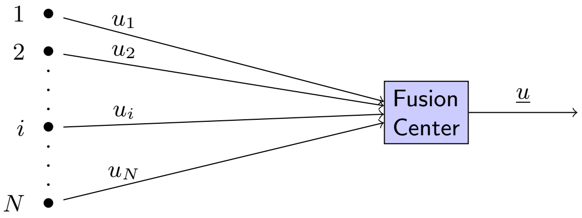


Outline

- Introduction to detection theory
- Detection from one sensor
  - Decision rules, MAP, LRT, ML
  - The Neyman-Pearson criterion
- Detection from multiple sensors
  - The Likelihood Ratio Test
  - The counting rule

Multiple sensors

Distributed detection method



- In each sensor $y_i(t) = \begin{cases} n_i(t) & \text{if } H_0 \text{ happened,} \\ \mu_i + n_i(t) & \text{if } H_1 \text{ happened.} \end{cases}$

  and a decision $u_i(t) = \begin{cases} 0 & \text{if } H_0 \text{ happened,} \\ 1 & \text{if } H_1 \text{ happened.} \end{cases}$ is taken

- The fusion center takes an overall decision $\underline{u} = f(u_1, ..., u_N)$ after collecting the decisions from each sensor

Multiple sensors

In case of multiple sensors, the definitions of probabilities are as follows

$$\mathrm{P}_{F_i} = \Pr\left(u_i = 1|H_0\right) = \int_{\gamma_i}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = Q\left(\gamma_i\right) \triangleq \Pr_i\left(H_1|H_0\right)$$

$$\mathrm{P}_{M_i} = \Pr\left(u_i = 0|H_1\right) = \int_{-\infty}^{\gamma_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\mu_i)^2}{2}} dt = 1 - Q\left(\gamma_i - \mu_i\right) \triangleq \Pr_i\left(H_0|H_1\right)$$

$$\mathrm{P}_{D_i} = \Pr\left(u_i = 1|H_1\right) = \int_{\gamma_i}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\mu_i)^2}{2}} dt = Q\left(\gamma_i - \mu_i\right) \triangleq \Pr_i\left(H_1|H_1\right)$$

where P_{F_i}, P_{M_i} and P_{D_i} the probabilities of false alarm, miss detection and detection of a particular node i respectively.

Decision function f

How to design an optimal decision function f?

- Option 1: The Likelihood Ratio Test
- Option 2: The counting decision rule

The Likelihood Ratio Test

By applying the Likelihood Ratio Test (LRT) we get

$$\frac{\Pr(\underline{u}|H_1)}{\Pr(\underline{u}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\Pr(H_0)}{\Pr(H_1)} \overset{ML}{\Rightarrow} \frac{\Pr(\underline{u}|H_1)}{\Pr(\underline{u}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

Using the Bayes' rule, ML criterion becomes

$$\frac{\Pr(H_1|\underline{u})}{\Pr(H_0|\underline{u})} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

The Likelihood Ratio Test

Theorem

$$\log \frac{\Pr\left(H_1 | \underline{u}\right)}{\Pr\left(H_0 | \underline{u}\right)} = \log \frac{\Pr\left(H_1\right)}{\Pr\left(H_0\right)} + \sum_{S^+} \log \frac{1 - \mathrm{P}_{M_i}}{\mathrm{P}_{F_i}} + \sum_{S^-} \log \frac{\mathrm{P}_{M_i}}{1 - \mathrm{P}_{F_i}}$$

where S^+ *the set of sensors that decide for 1*
S^- *the set of sensors that decide for 0*

The Likelihood Ratio Test

<u>Proof</u>:

We have

$$\Pr(H_1|\underline{u}) = \frac{\Pr(H_1,\underline{u})}{\Pr(\underline{u})} = \frac{\Pr(H_1)}{\Pr(\underline{u})} \prod_{S_+} \Pr(u_i = +1|H_1) \prod_{S_-} \Pr(u_i = 0|H_1)$$

$$= \frac{\Pr(H_1)}{\Pr(\underline{u})} \prod_{S_+} (1 - P_{M_i}) \prod_{S_-} P_{M_i} \,.$$

In a similar manner,

$$\Pr(H_0|\underline{u}) = \frac{\Pr(H_0)}{\Pr(\underline{u})} \prod_{S_+} (1 - P_{F_i}) \prod_{S_-} P_{F_i} \,.$$

Thus, we have that

$$\log \frac{\Pr(H_1|\underline{u})}{\Pr(H_0|\underline{u})} = \log \frac{\Pr(H_1)}{\Pr(H_0)} + \sum_{S_+} \log \frac{1 - P_{M_i}}{P_{F_i}} + \sum_{S_-} \log \frac{P_{M_i}}{1 - P_{F_i}} \,.$$

The counting rule

Another decision criterion is the counting rule

$$\Lambda = \sum_{i=1}^{N} u_i \underset{H_0}{\overset{H_1}{\gtrless}} \Gamma$$

where Γ is the decision threshold

Assuming that the decision threshold, γ_i, is the same (γ) for all sensors, the probability of false alarm is now

$$\Pr(\Lambda \geq \Gamma | N, H_0) = \sum_{i=\Gamma}^{N} \left(\begin{array}{c} N \\ i \end{array} \right) \mathrm{P}_{F_i}^{i} (1 - P_{F_i})^{N-i}$$

Applying the Laplace - de Moivre approximation,

$$\Pr(\Lambda \geq \Gamma | N, H_0) \simeq Q\left(\frac{\Gamma - N \cdot P_{F_i}}{\sqrt{N \cdot P_{F_i} (1 - P_{F_i})}} \right)$$

Summary

We have studied the basic principles regarding the detection from one or multiple sensor(s).

We have also seen certain representative decision rules for detecting events out of uncertain (noisy) observations

Next lecture

Next lecture, we study how to perform static estimation from noisy measurements of the sensors