Cyber-Physical Security in Time-Critical Systems (EL2850)

Homework 1-3 (Module 3, updated 251006)

Due by: October 13, at 19:00

Pass criterion: 70% (15/22 points)

Instructions: Individually submit your solutions in PDF format in Canvas. Preferably, write your answers using LaTeX. Describe all steps in your solution. Include Python/MATLAB code if it has been used. Discussing the problems with fellow students is allowed, but the answers should explicitly state any collaboration.

1. False alarms and missed detection probabilities

This exercise aims to understand the concepts of false alarms and missed detection. Consider a residual signal r modeled as the outcome of a random variable

$$r = f + v,$$

where f is the deterministic fault signal we want to detect, and v is a normally distributed random variable (modeling measurement noise) with mean $\mu = 0$ and standard deviation $\sigma = 1$. We consider the following two hypotheses:

$$\begin{aligned} H_0 &: \quad f = 0, && \text{no fault (null hypothesis)}, \\ H_1 &: \quad f \neq 0 && \text{fault present (alternative hypothesis)}. \end{aligned}$$

Based on a residual sample r, a detector test is defined that generates an alarm when $|r| > \gamma$, where $\gamma > 0$ is a predetermined threshold. The outcomes of the test are defined by

$$\begin{aligned} D_0 &: \quad |r| \leq \gamma && \text{Do not reject } H_0 \text{ (no alarm generated).} \\ D_1 &: \quad |r| > \gamma && \text{Reject } H_0 \text{ (alarm generated).} \end{aligned}$$

(a) Sketch residual distribution when $f = 0$ and $f = 3$ and mark a suitable threshold.

(b) Define the false alarm probability α. The false alarm probability corresponds to an area in the sketch from problem 1a. Mark this area.

(c) If Z is a normally distributed random variable with mean $\mu = 0$ and standard deviation $\sigma = 1$ then let $\Phi(z)$ be the cumulative distribution function for Z, i.e.

$$\Phi(z) = \Pr(Z \leq z) = \int_{-\infty}^{z} f_Z(z)\, dz$$

where $f_Z(z)$ is the known normal distribution curve

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

Using $\Phi(z)$, sketch the probability for false alarms as a function of the threshold γ.

(d) Define the probability for missed detection β given a fault with the size $f = f^1 \neq 0$. Like the false alarm probability, the probability for missed detection corresponds to an area in the sketches from problem 1a. Mark this area.

(e) Using $\Phi(z)$, sketch the probability for missed detection given a fault with the size $f = 3$ as a function of the threshold γ.

(f) What is the best possible false alarm probability and the probability for a missed detection at a test?

(10p)

2. Detection performance

Consider the same residual as in Problem 1 and let the standard deviation be $\sigma = 2$.

(a) Calculate a numerical value on the threshold γ such that the probability of false alarm becomes $\alpha = 0.01$.

(b) The power function is a way to specify detector performance and is defined as

$$\beta_{\text{pow}}(f^1) = \Pr(|r| > \gamma \mid f = f^1).$$

The power function thus describes the probability of detection for different fault sizes. Calculate the power function for the calculated threshold value in Problem 2a. Verify using the power function that the probability of false alarm is $\alpha = 0.01$. Plot the power function in Python/ MATLAB.

(c) For a given fault size, for example, $f = 5$, it can be interesting to study how the threshold influences the compromise between achieving a low probability of false alarm α, and a high probability of detection $1 - \beta$. Investigate the compromise by drawing $1 - \beta$ as a function of α and interpret the results. This curve is called the ROC (Receiver Operating Characteristic) curve.

How does the curve change if f varies?

(5p)

3. CUSUM

The CUSUM algorithm is formally derived under the assumption that the distribution of the residual is known, both in the fault-free case and when a fault has occurred. This is often not a realistic situation, and this exercise tries to illustrate how the CUSUM algorithm can be useful even if detailed statistical knowledge is missing.

The file named `residualdata.mat` contains a residual signal over 1000 time steps where a fault occurs at some point. Import the data into your workspace. In Python, use `import scipy.io` and `scipy.io.loadmat('residualdata.mat')`. In MATLAB, use `load('residualdata.mat')`. This should create a variable r in your workspace, with 1 row and 1000 columns. The first column contains $r[0]$, the second contains $r[1]$, and so forth.

(a) From the data in r, estimate the parameters required to run a CUSUM algorithm.

(b) Use the CUSUM algorithm to, in case of an alarm, estimate the time the fault occurred.

(7p)