The 4th IET International Conference on Railway Condition Monitoring
RCM 2008

Switch Mechanism Diagnosis using a Pattern Recognition Approach

F. Chamroukhi, A. Samé, P. Aknin

The French National Institute for Transport and Safety Research (INRETS-LTN)

M. Antoni

The French National Railway Company (SNCF)

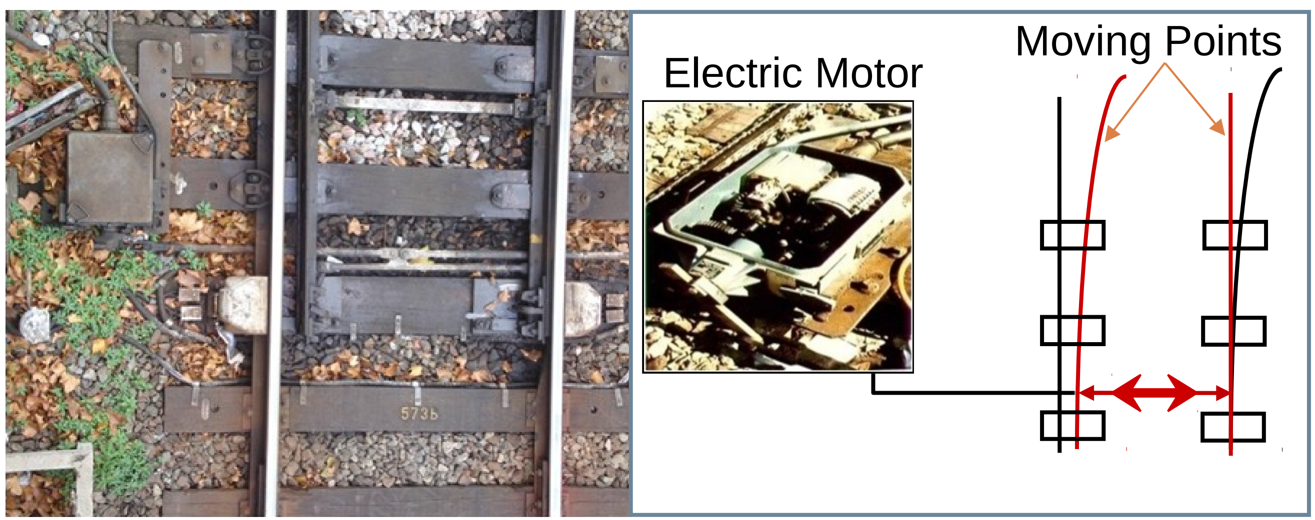
19 June 2008, Derby, UK

Overview of the presentation

- Context
- The proposed pattern recognition approach
  - General principle of a pattern recognition approach
  - Signals parameterization
  - Parameters learning: Mixture Distribution Modeling
  - Signals classification
- Experimental study
- Conclusion and future works

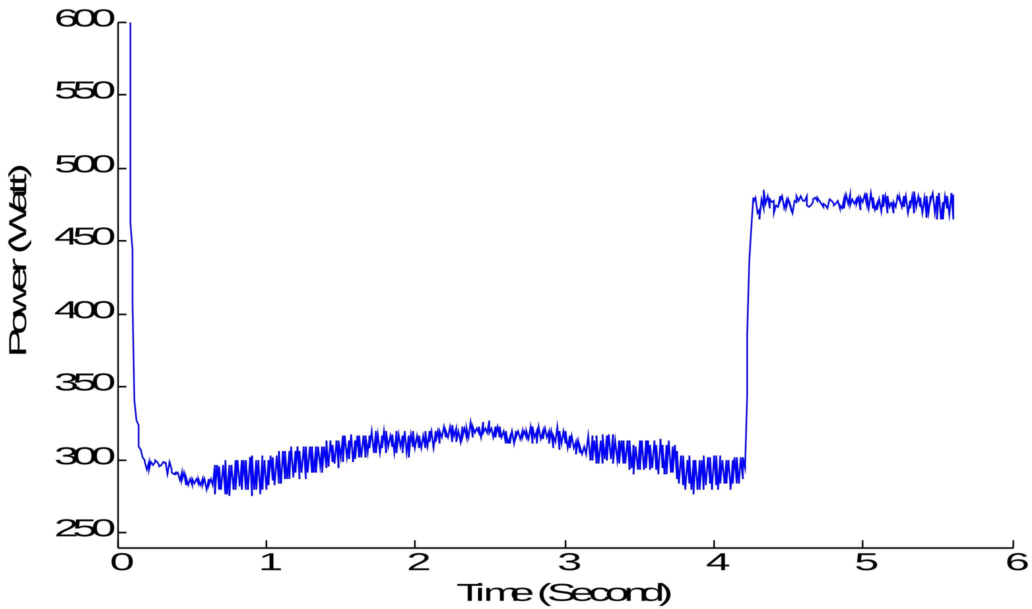
Context

- Switch mechanism diagnosis
- Considered switches:
  - Operated by an electric motor
  - Equiped with a Clamp-Lock system (« Verrou-Carter-Coussinet » in french)



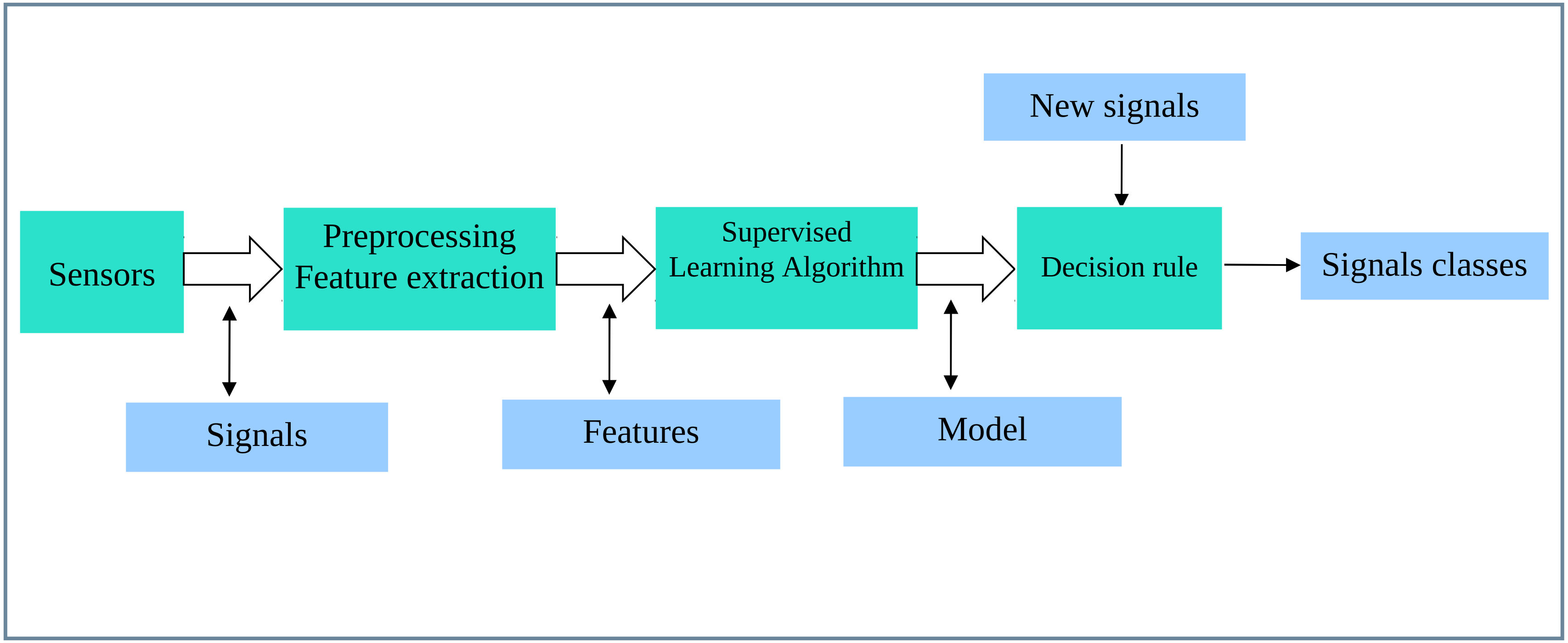
Acquired signals

- Measurements of the electrical power consumption during the switch actuation period
- Sampling frequency: 100 hz
- Length of each signal: 550 points



The proposed pattern recognition approach

General principle of a pattern recogniton approach

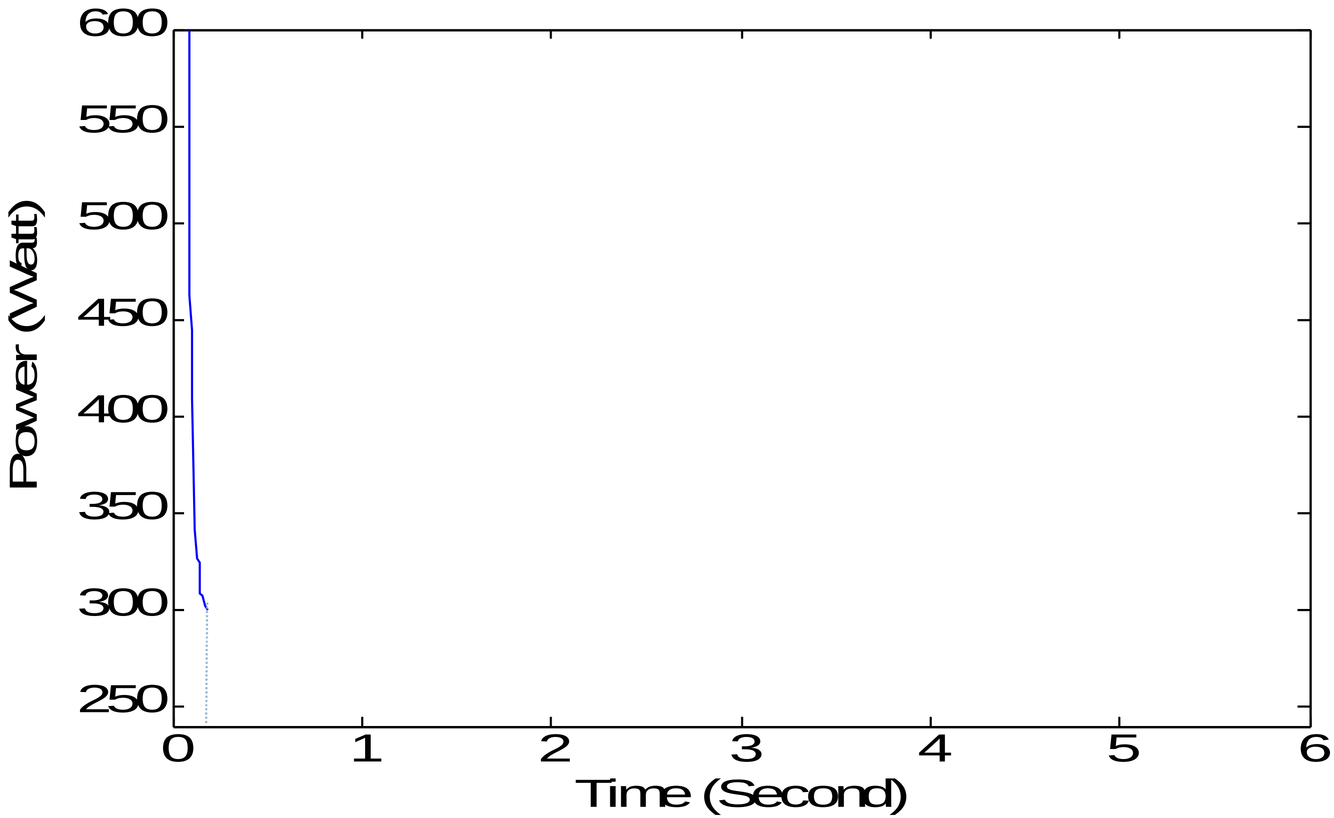


Signals parameterization: Feature extraction

- The switch actuation consists of successive mechanical motions of different parts of the mechanism:
  - starting phase
  - points unlocking
  - points translation
  - points locking
  - friction phase
- These motions are observed on the shape of the signal

The different phases of a switch actuation

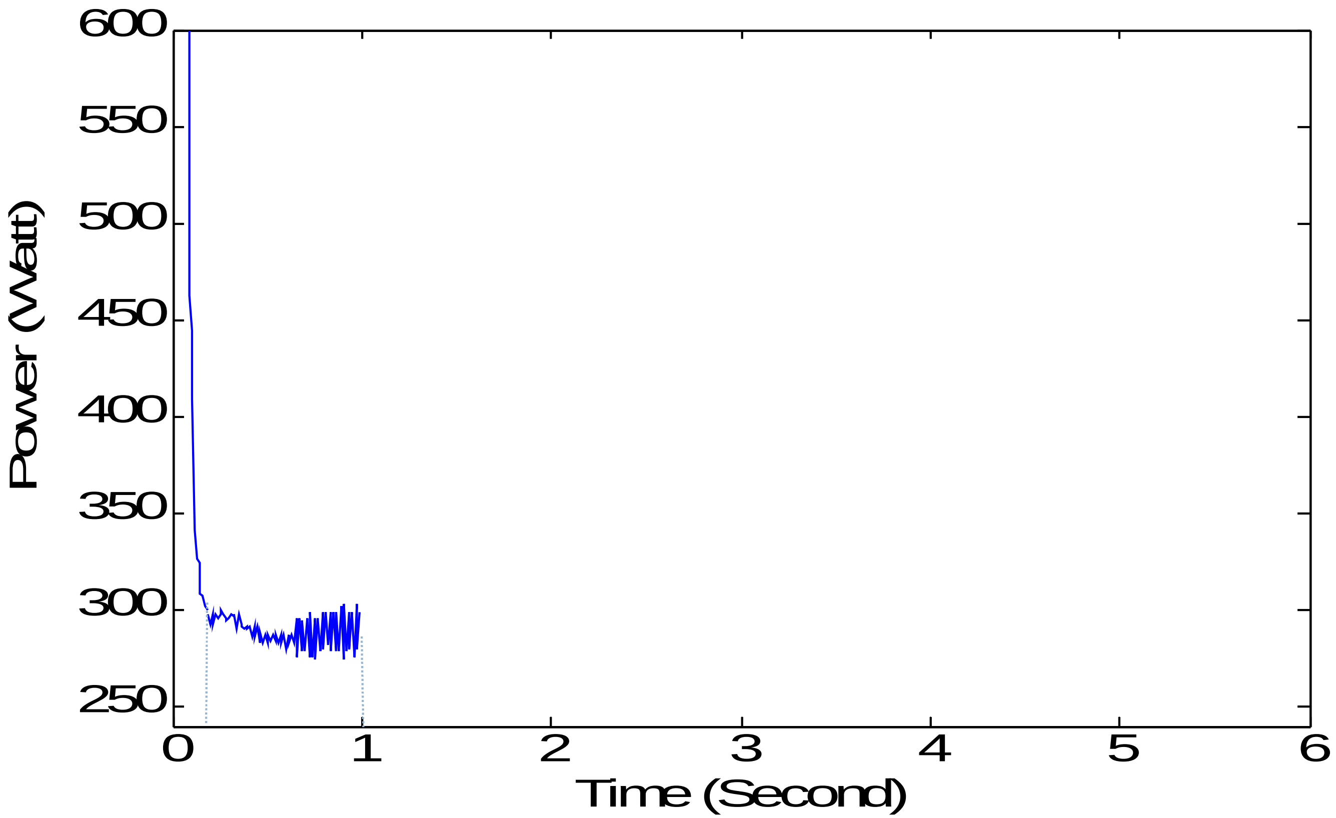
- The starting phase



Starting

The different phases of a switch actuation

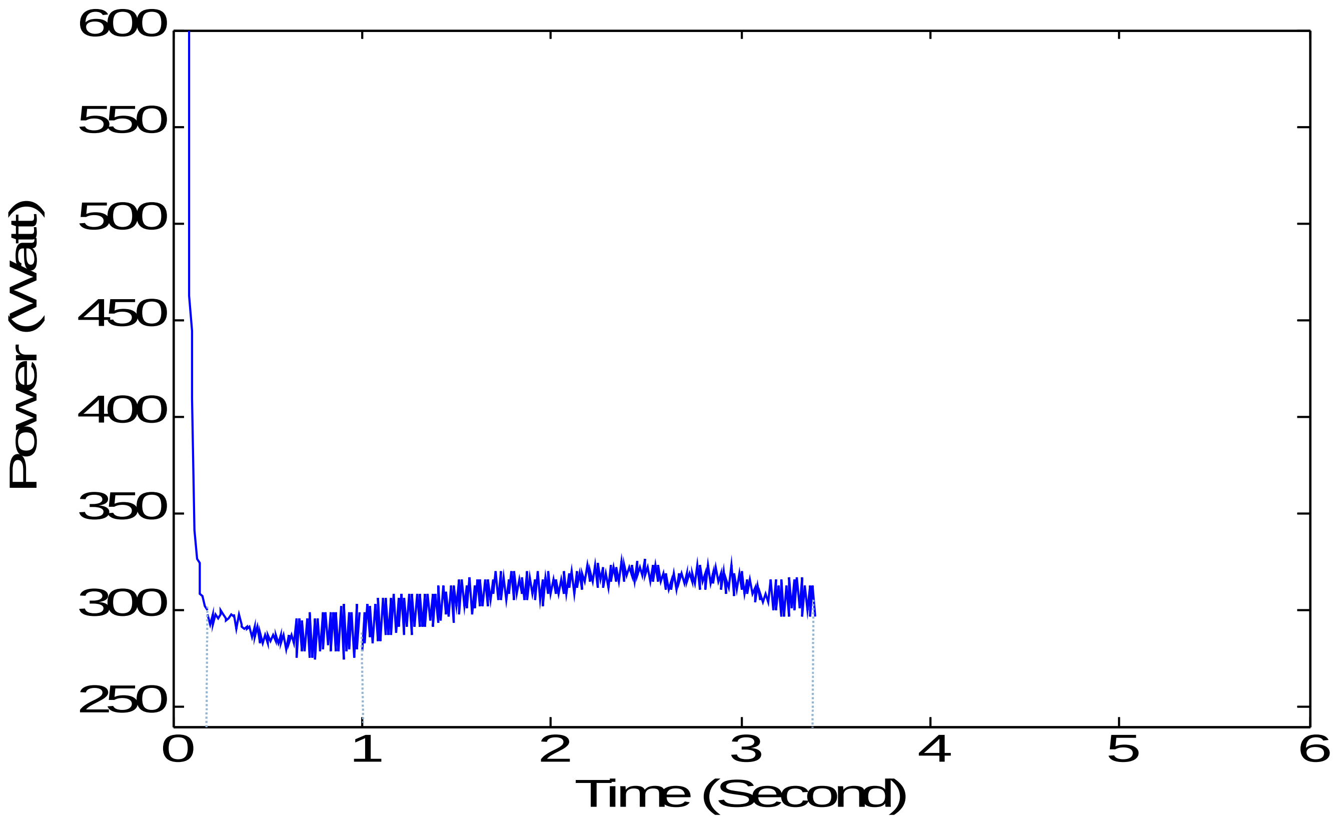
- The unlocking phase



Starting Unlocking

The different phases of a switch actuation

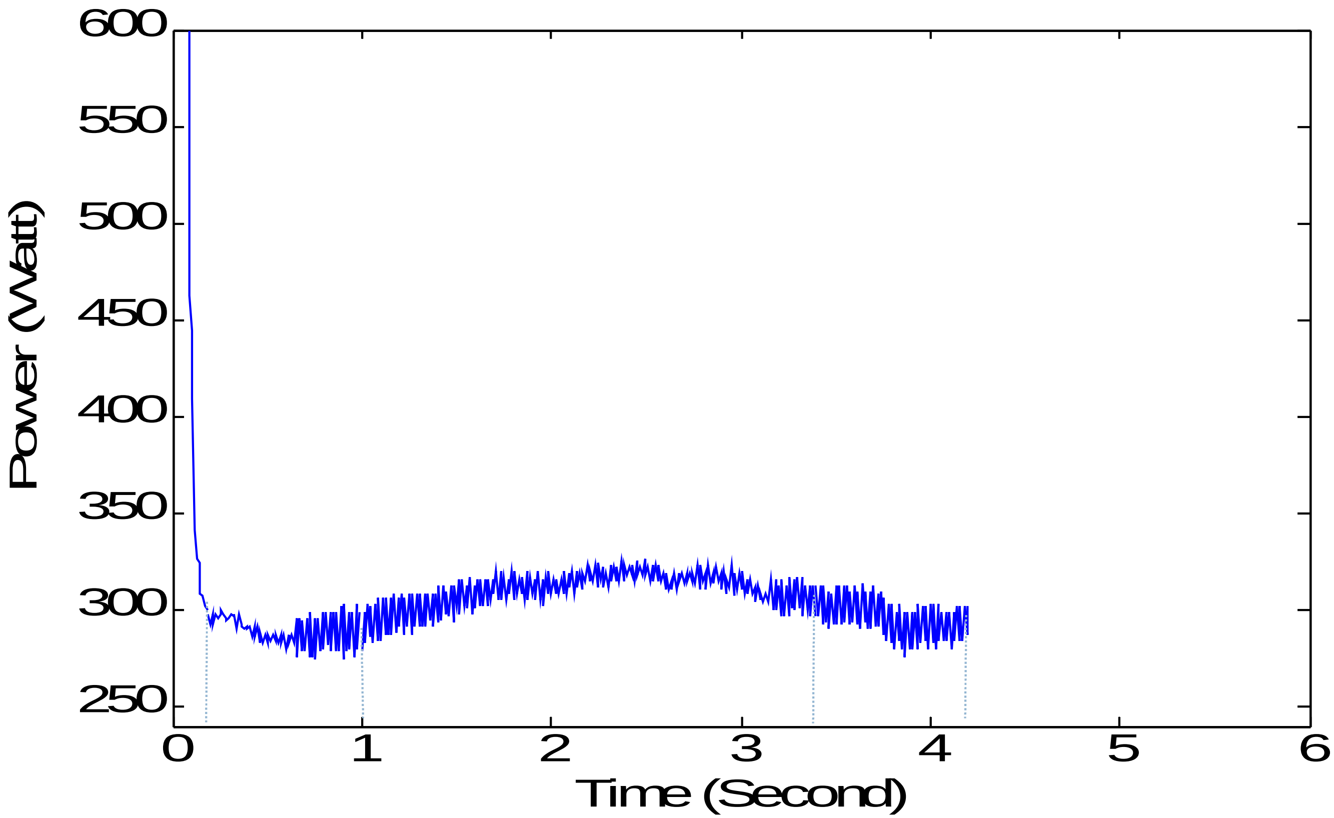
- The translation phase



Starting Unlocking Translation

The different phases of a switch actuation

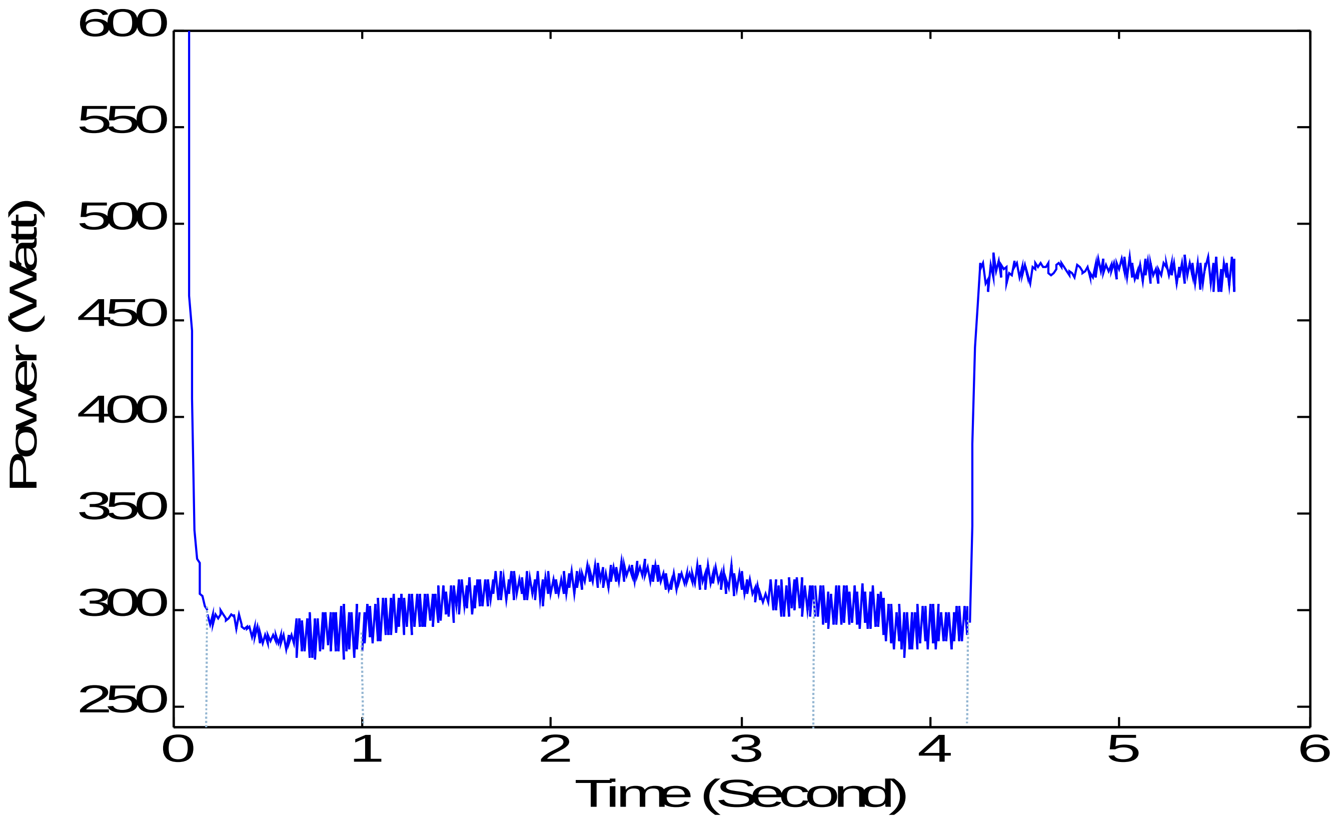
- The locking phase



Starting Unlocking Translation Locking

The different phases of a switch actuation

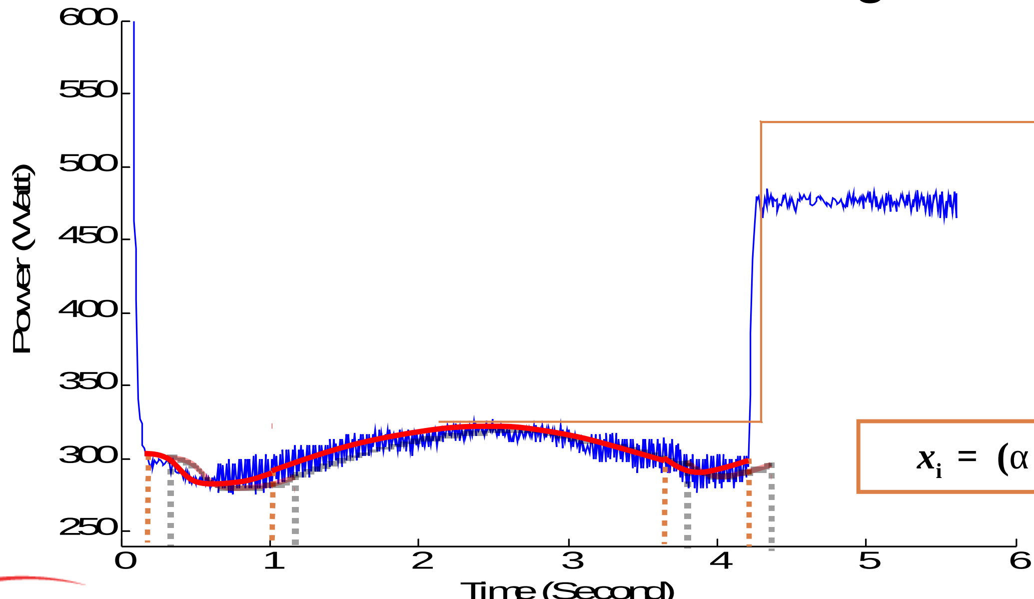
- The friction phase



Starting Unlocking Translation Locking Friction

Feature extraction

- Each signal is described by the set of the features of its three main phases
  - Polynomial fitting
  - Parameter vector: polynomial coefficients, min, max, mean and variance of the signal in each segment



$$a_0 + a_1 t + \ldots + a_p t^p$$

$$\alpha_i = (a_0, a_1, \ldots, a_p)$$

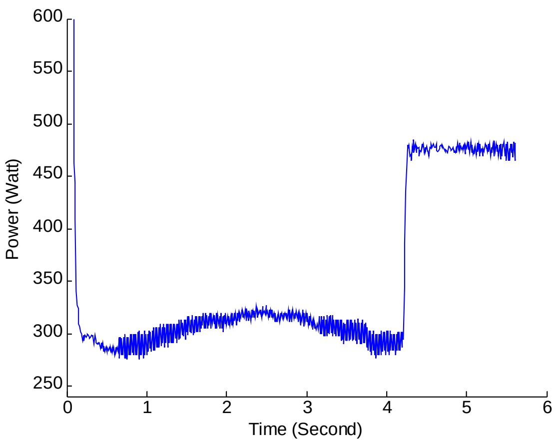
$$x_i = (\alpha_1 ; \alpha_2 ; \alpha_3 ; \text{min} ; \text{max} ; \text{mean} ; \text{variances})$$

→ For each signal we have 21 parameters instead of 550 points!

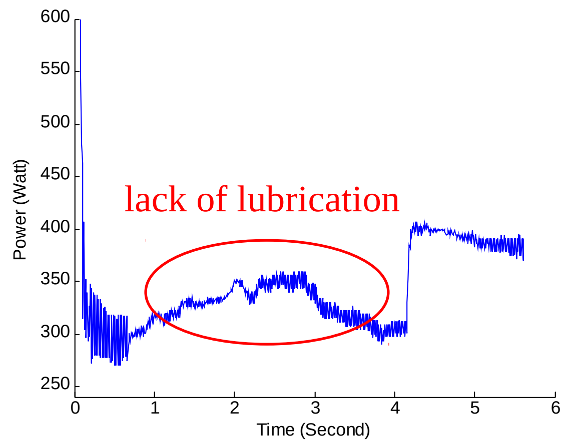
Learning parameters

- The three considered classes

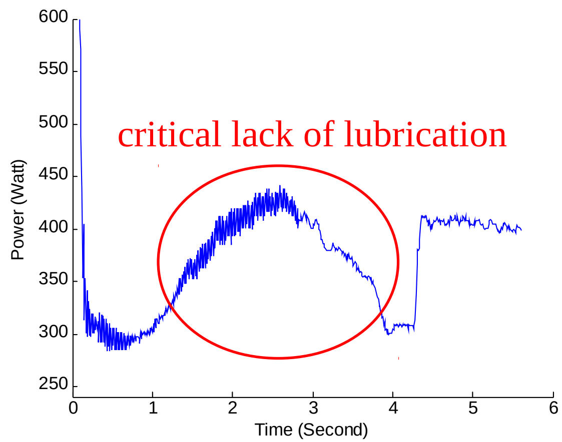
C_1: class without defect



C_2: class with minor defect



C_3: class with critical defect



Learning parameters: Mixture Discriminant Analysis (MDA)

Why Mixture Discriminant Analysis ?

- In classical Linear Discriminant Analysis (LDA), each class is modeled by a single Gaussian density
- For complex classes, a single density is insufficient
- ► Proposed solution: Gaussian Mixture Density
- MDA is a probabilistic discimination method based on Gaussian Mixture Model (GMM)

Advantages:

- MDA allows to model classes more precisely
- Improve the correct classification rate

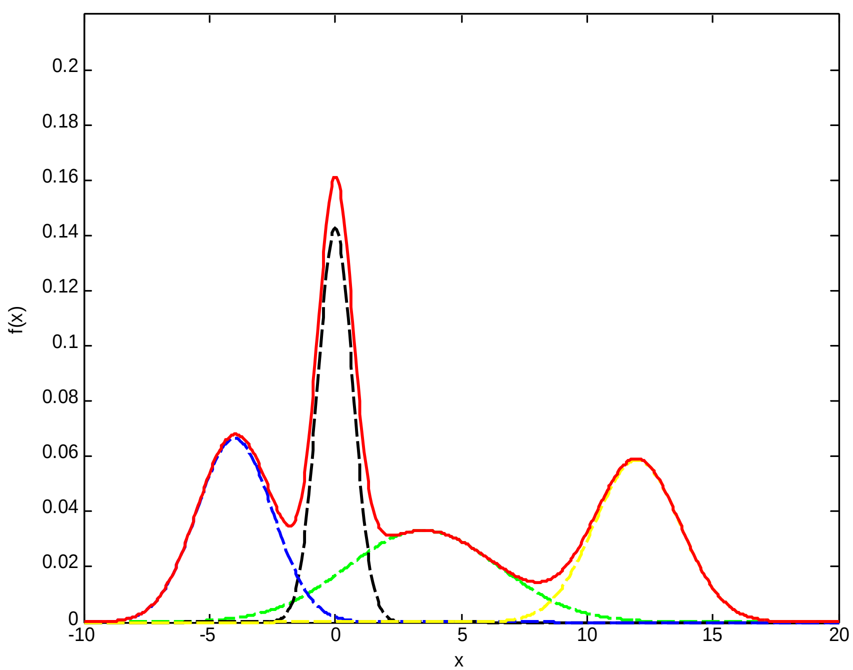
Gaussian Mixture Model (GMM)

- The mixture density for class C_k :

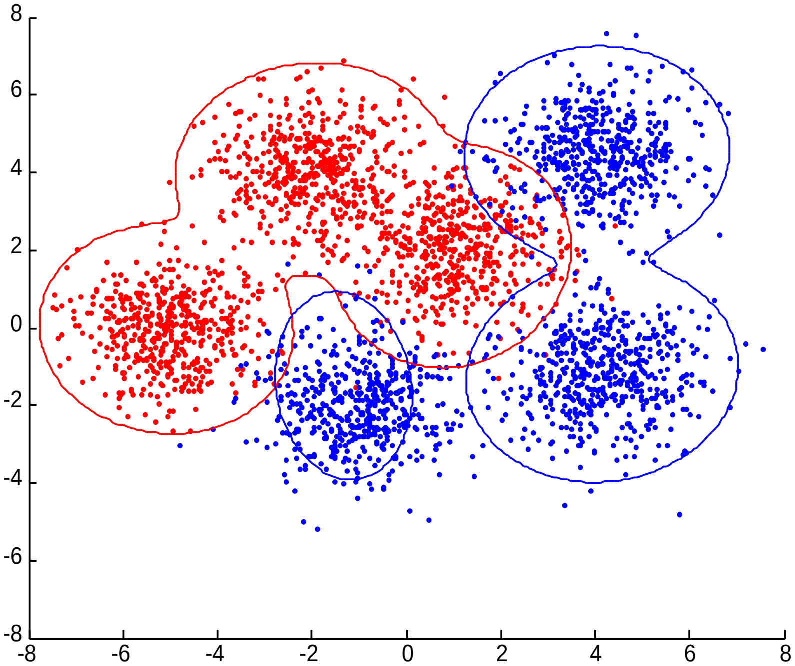
$$f(x_i|C_k) = \sum_{r=1}^{R_k} \pi_r \mathcal{N}(x_i|m_r, \Sigma_r)$$

  - x_i is the feature vector extracted from the ith signal
  - R_k is the number of densities of the mixture
  - The proportions of the mixture verify $\sum_{r=1}^{R_k} \pi_r = 1$
  - $\mathcal{N}(x_i|m_r, \Sigma_r)$ is the Gaussian probability density function with mean m_r and covariance matrix Σ_r
- $\theta_k = (\pi_1, \ldots, \pi_{R_k}, m_1, \ldots, m_{R_k}, \Sigma_1, \ldots, \Sigma_{R_k})$: parameter vector of the class C_k to be estimated.

Examples of GMM distributions



A GMM density in dimension 1



A bidimensional data set simulated according to two GMM distributions

Estimation of the mixture model parameters

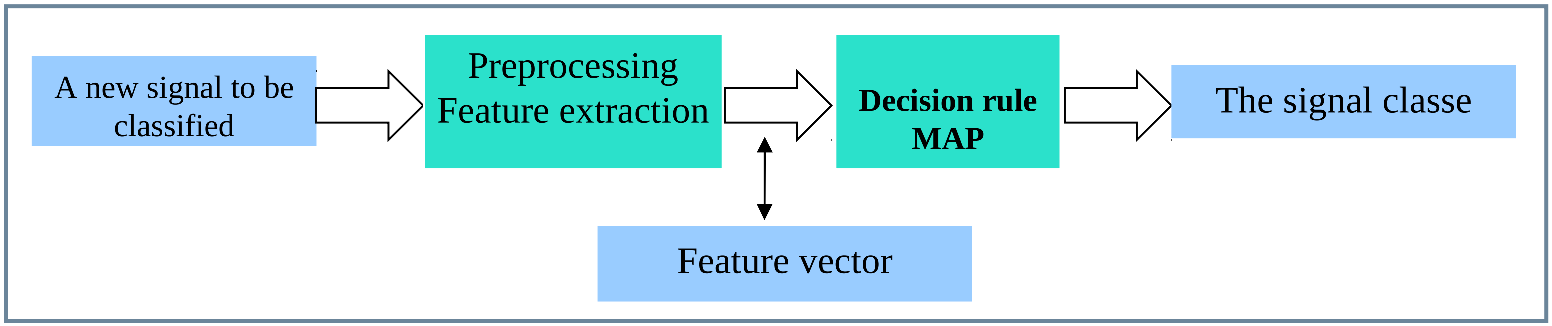
- Maximum Likelihood method
  - Log-likelihood:

$$L(\theta_k) = log \prod_{i=1}^{n} f(x_i|C_k) = \sum_{i=1}^{n} log\Big[\sum_{r=1}^{R_k} \pi_r \mathcal{N}(x_i|m_r, \Sigma_r)\Big]$$

  - The maximization is performed by a specific algorithm: the Expectation-Maximization (EM) algorithm
  - The optimal number of Gaussian distributions R_k for each class is computed by maximizing the Bayesian Information Criterion (BIC)

How to classify signals ?

- Use the Maximum A Posteriori (MAP) rule



- Assign each signal represented by x_i to the class k* which maximizes the posterior probabilities

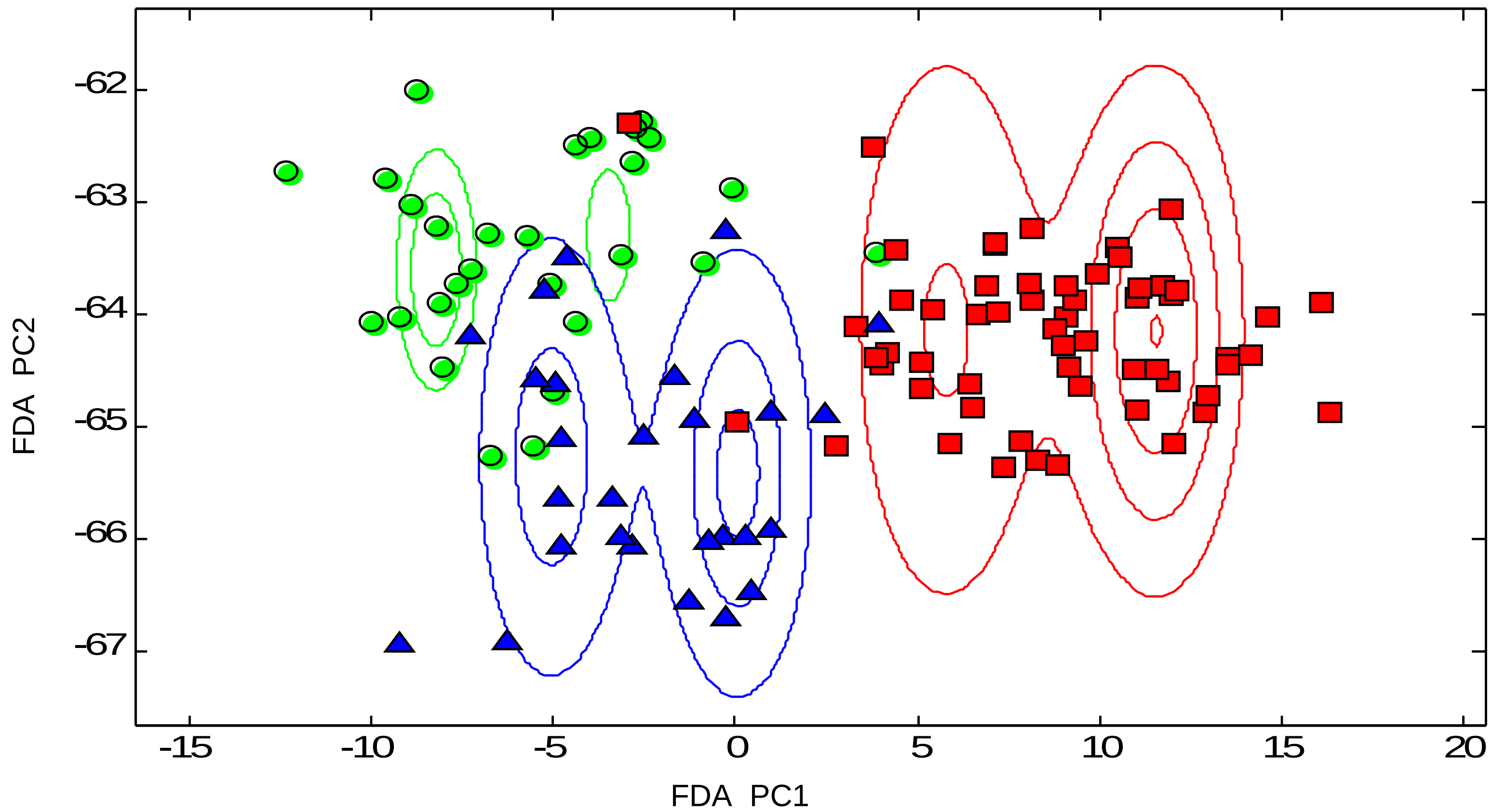
$$k^* = \underset{k}{argmax}\ P(C_k|x_i) \ \text{ where } \ P(C_k|x_i) = \frac{f(x_i|C_k)P(C_k)}{\sum_{\ell=1}^{K} f(x_i|C_\ell)P(C_\ell)}$$

Experimental study

- Database: 119 labellized signals
  - 90 Signals used for learning (supervised learning)
  - 29 signals used to evaluate the classifier

- Comparison to alternative classification approaches
  - Neural Network (Based on Multilayer Perceptron)
  - K-Nearest Neighbours
  - Bayesian discrimination approach (A single gaussian density for each class)

Estimated GMM distributions into the principal factor discriminant plane



Class1 = green circle, class 2 = blue triangle, class 3 = red square

Results

- The correct classification rate obtained with the four methods :

Approach	Correct Classification Rate
MDA	**95 %**
NN	90 %
KNN	88 %
Bayesian disc. with one Gaussian	75 %

- The number of selected mixture components

Class	C1	C2	C3
Number of mixture components	4	2	2

Conclusion

- Development of a classification method based on Mixture Discriminant Analysis (MDA) in a switch mechanism diagnosis context
- This type of approach can be applied to various switch mechanisms since it simply requires the electric power consumption signals
- The experimental study on real signals has revealed some good performances of our approach, compared to alternative methods

Future Works

- Time monitoring of the state point over a sequence of actuations
- Envisaged approaches:
  - Regressive and autoregressive mixtures models
  - Hidden Markov Models