FAULT TEMPLATE EXTRACTION FROM INDUSTRIAL ALARM FLOODS

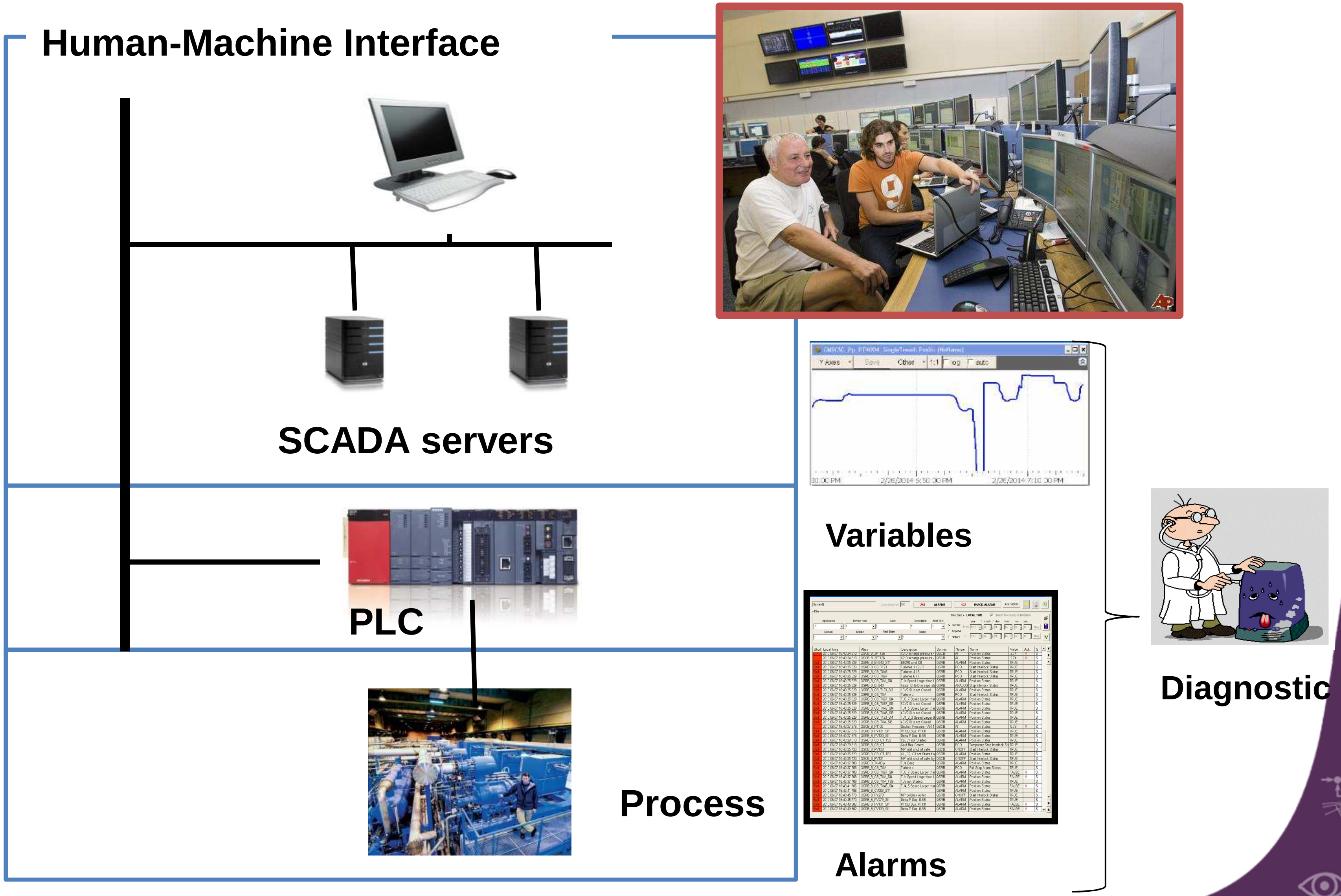
Sylvie Charbonnier, Nabil Bouchair, Philippe Gayet

gipsa-lab

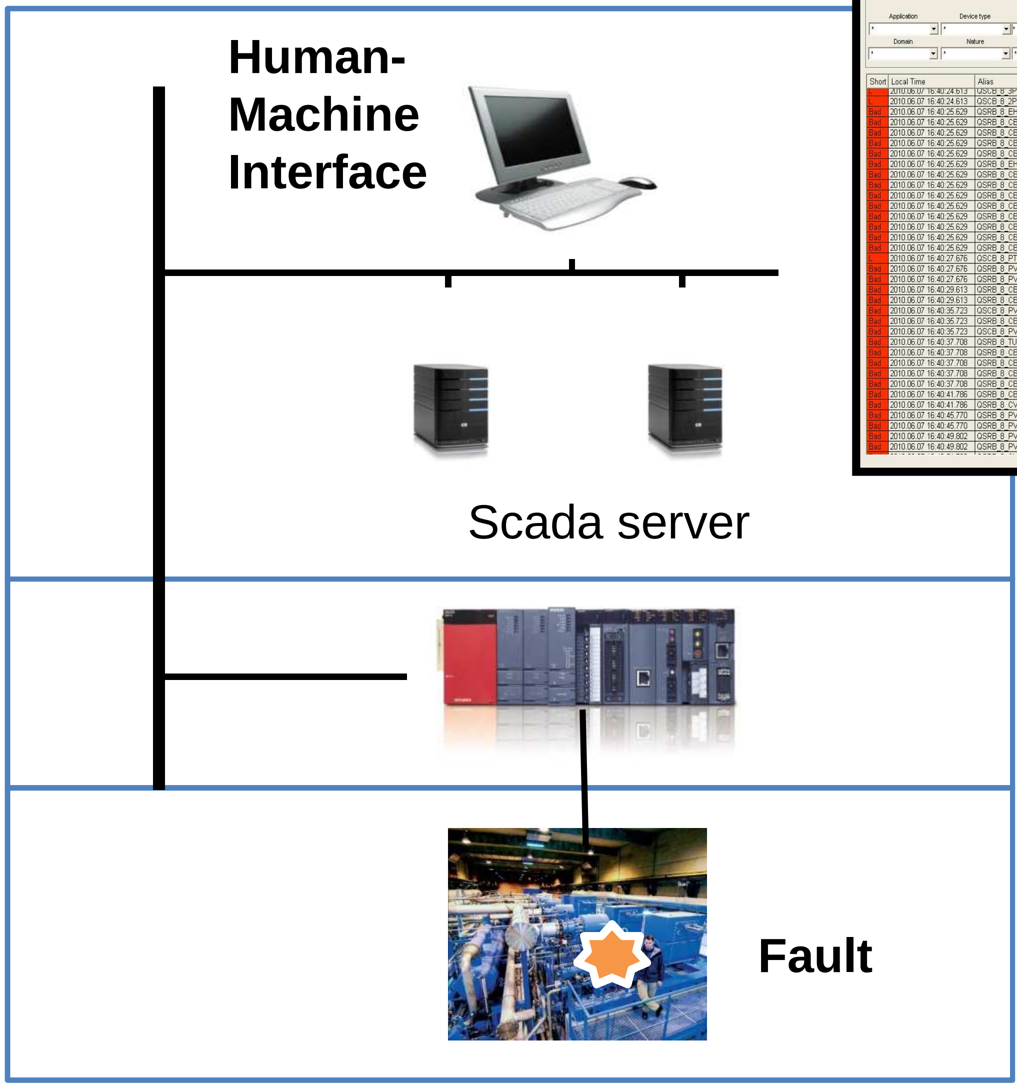
Grenoble | images | parole | signal | automatique | laboratoire

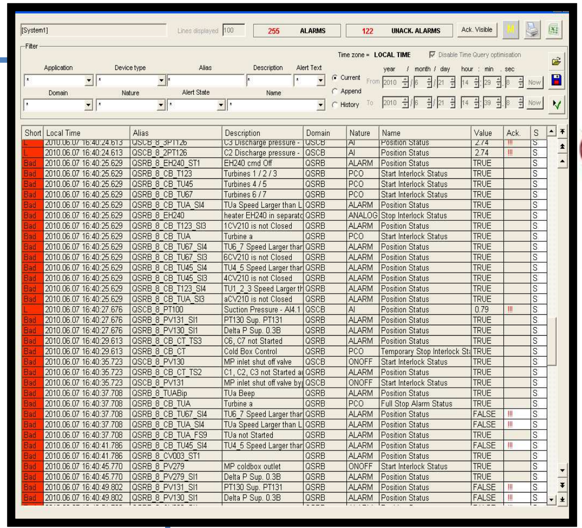
UNIVERSITÉ DE GRENOBLE

Industrial control systems based on SCADA architecture



Diagnostic may be complex

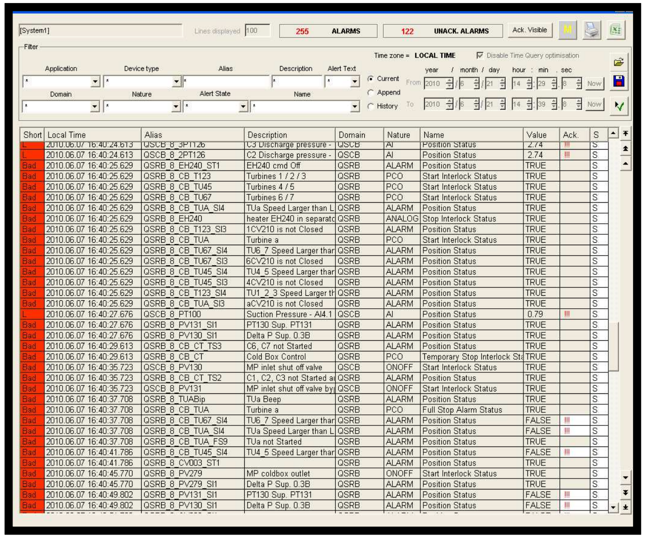




Alarm flood

More than 10 alarms during 10 min (ANSI/ISA-18,2)

To assist the operator confronted to an alarm flood



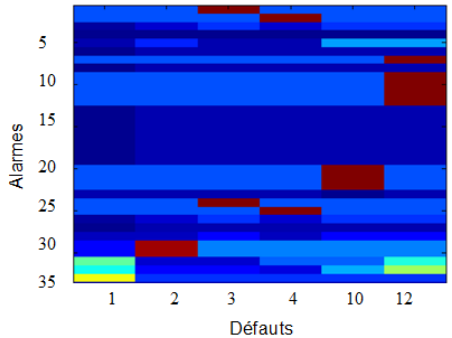
Fault template: learnt from a set of alarm lists recorded during the occurrence of the fault

- **sequential :**

The alarm order of appearence is used

- **vectorial**

Alarms are weighted according to their relevance to the fault



Learn information on the alarm generation process (off line)

- Which alarm is relevant?
- Which alarms are specific to a given fault?
- Is a fault diagnosable from the alarm list?
- Is there a specific pattern in the alarm flood ?

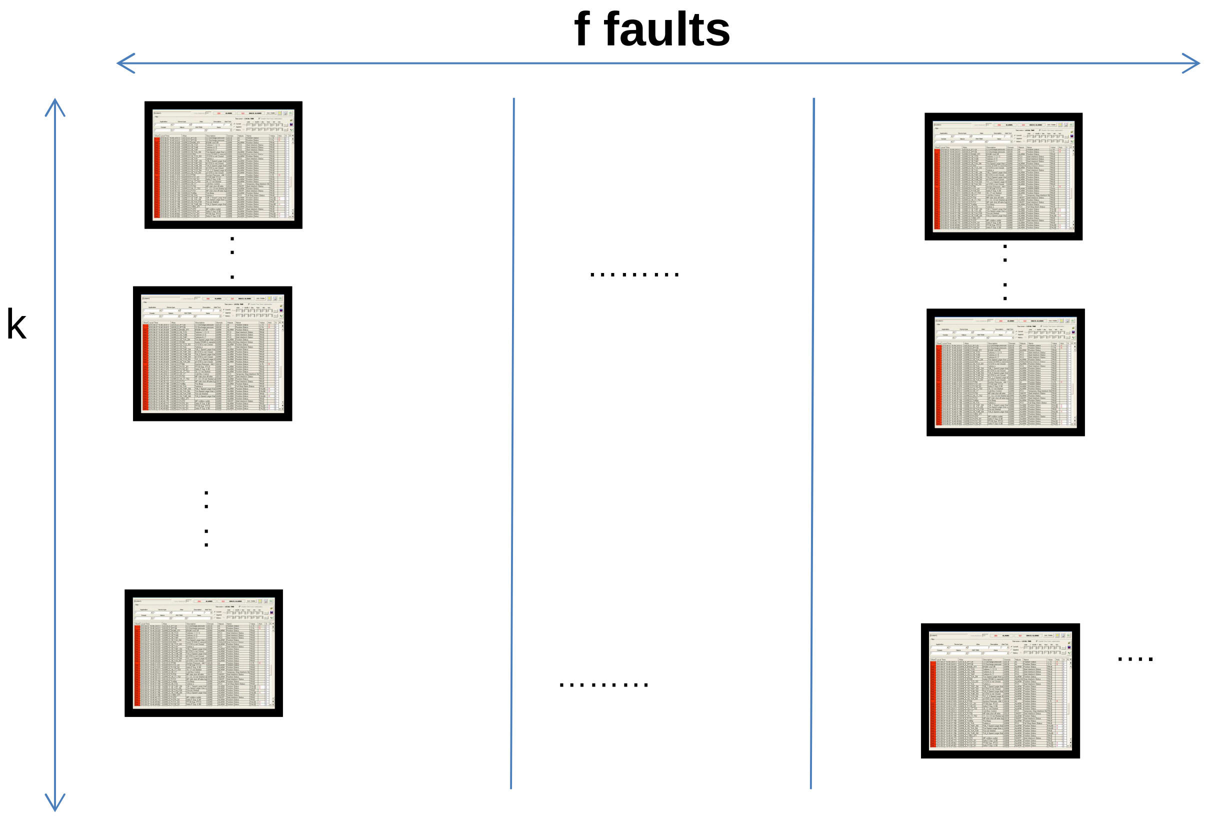
Decision support in operation (on line)

- Suggest possible faults

Outline

1. Methods
   A. Template extraction
      - sequential template
      - vectorial template
      - weighted sequential template
   B. Knowledge extraction and decision support
2. CERN LHC process
3. Results

Learning set



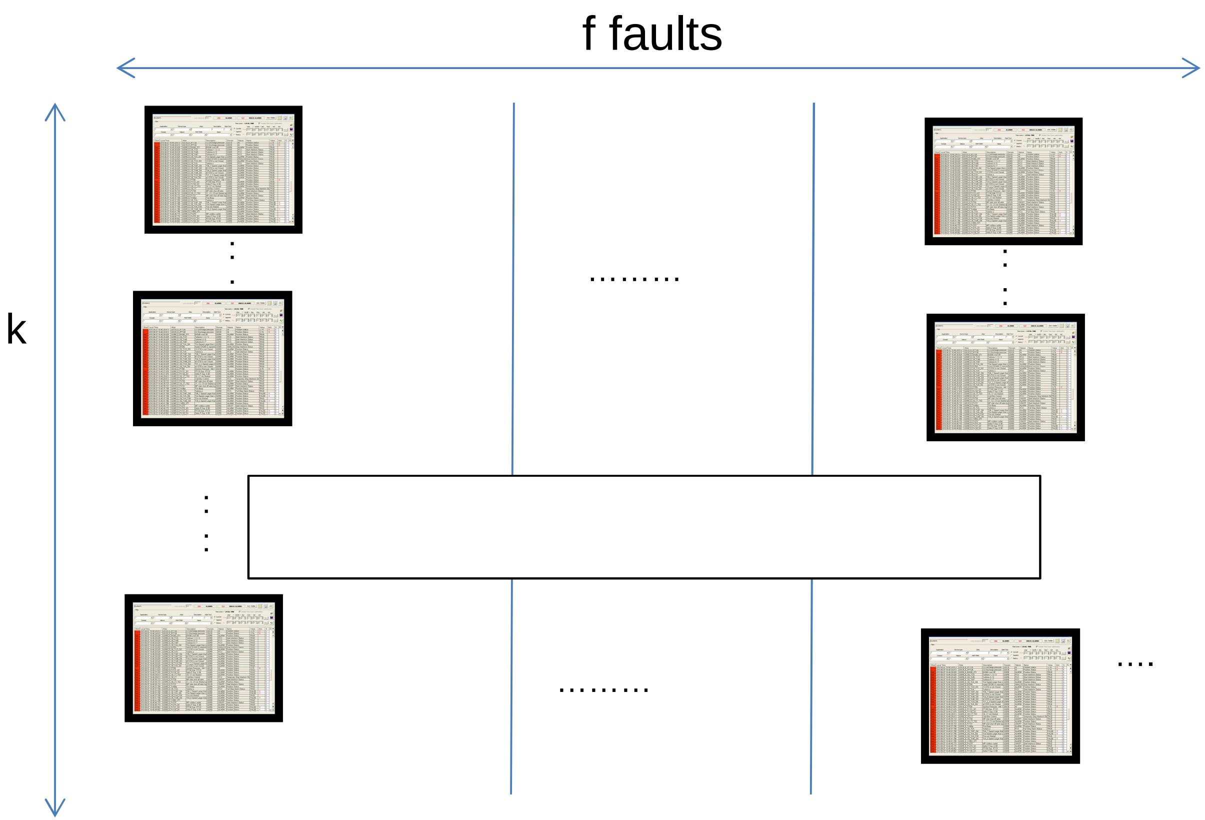
N = full number of alarms triggered by the control system

Sequential fault template

N = full number of alarms triggered by the control system

of size *N*

An alarm list = a symbolic sequence

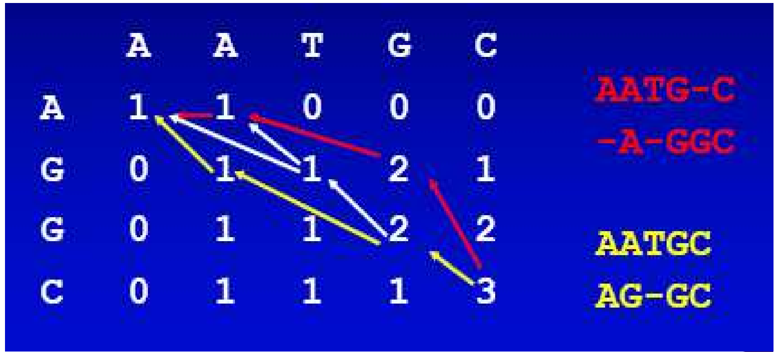


Sequential fault template : the minimal **sequence formed of alarms present in the same order on more than half the lists generated by fault n°q**

Template extraction using the Needleman and Wunsch algorithm

The **Needleman and Wunsch** algorithm (1970)

- Used in bio-informatic to compare sequences of genes
- Use dynamic programming



Global optimal alignment of 2 symbolic sequences – Insert gaps to increase similarity

Similarity between 2 sequences

Tuning parameters:

cost of a gap =0

C ( ) cost of susbtituting one element with another

+1 -1

$$C = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$

→ Similarity index:

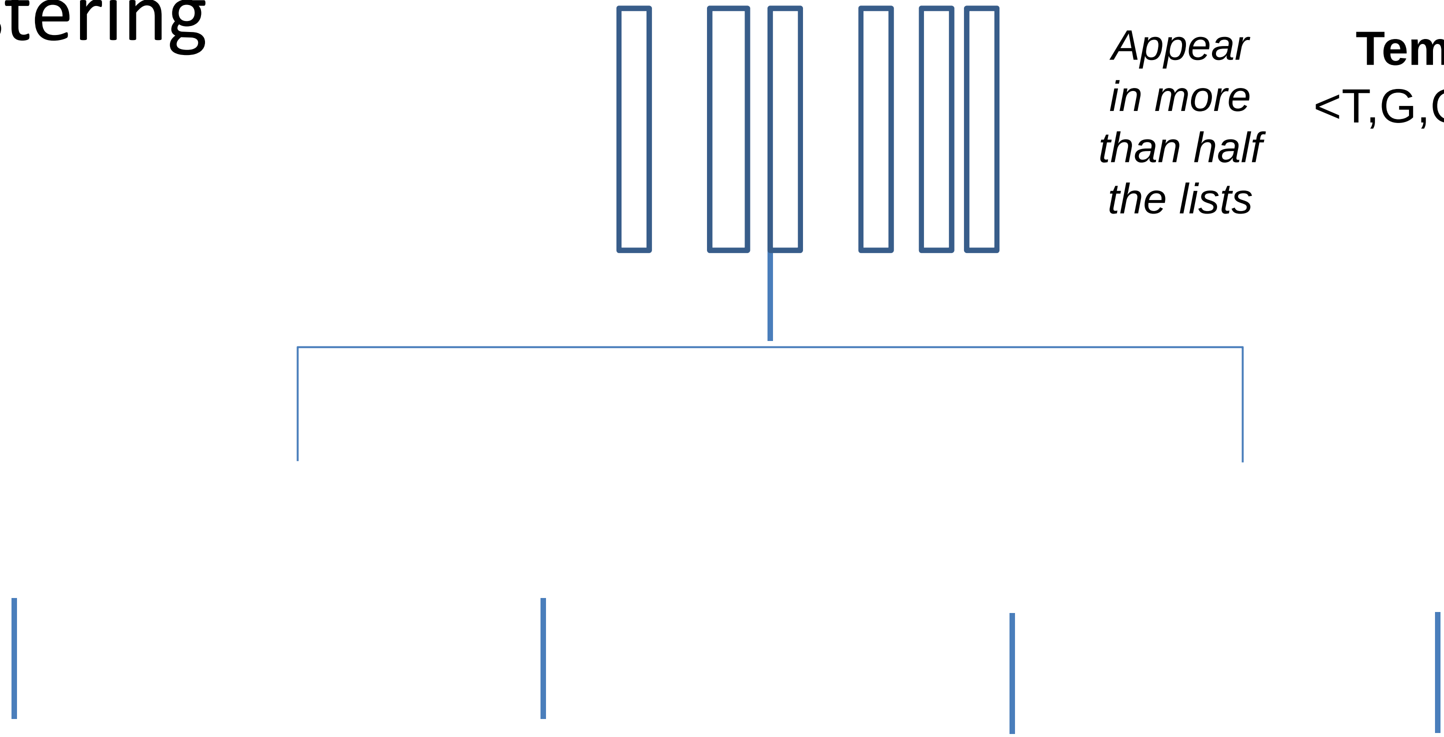
$$sim_{NWA}(S^{a^1}, S^{a^2}) = \frac{\sum_{p=1}^{l} \sigma(s_p^{a^1}, s_p^{a^2})}{l}$$

l

$\sigma(s_p^{a^1}, s_p^{a^2})$

Template extraction : multiple alignment

k sequences to align → ascendent hierarchical clustering



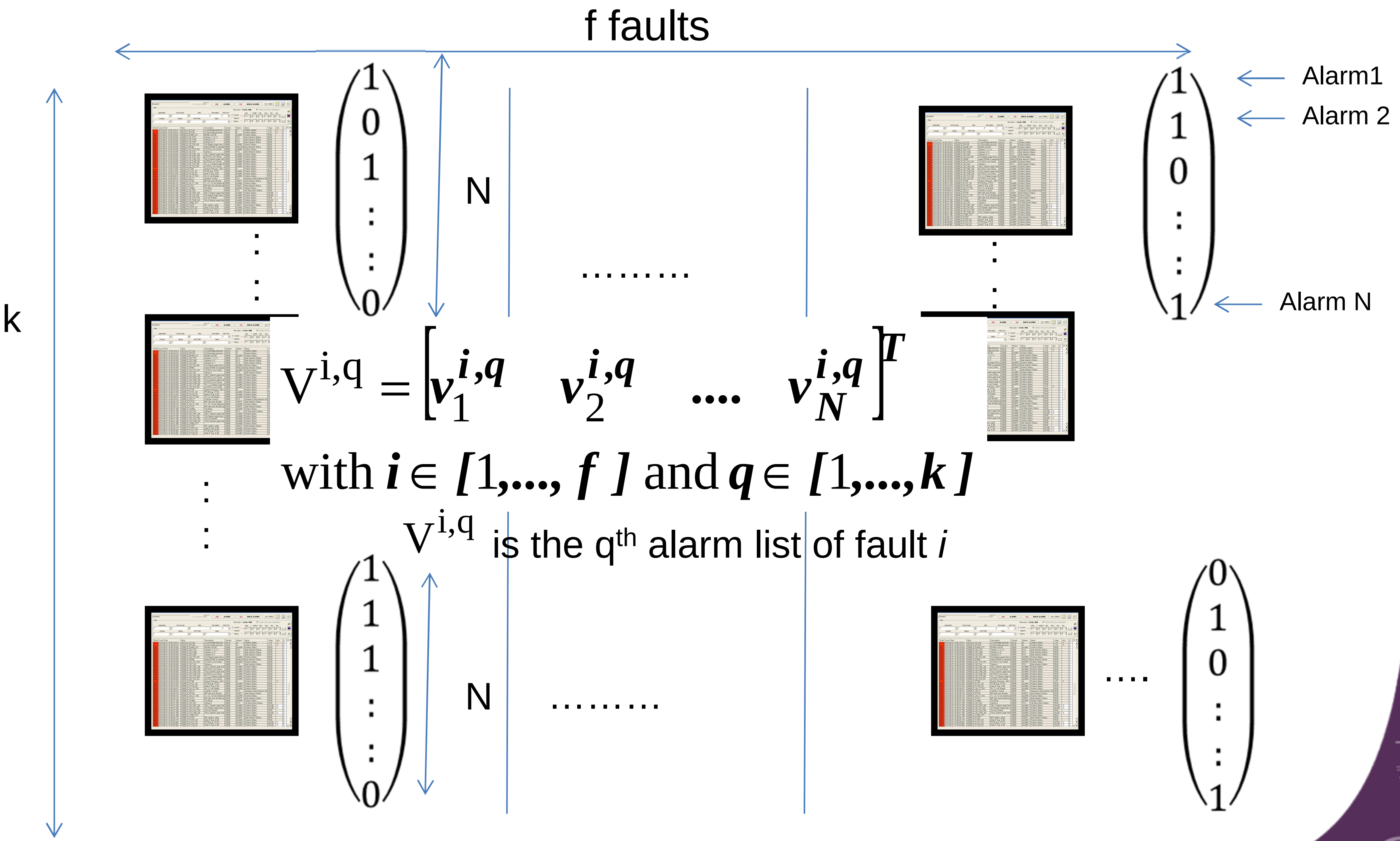
Appear in more than half the lists

Template :
<T,G,C,C,G,T>

Outline

1. Method
   A. Template extraction
      - sequential template
      - vectorial template
      - weighted sequential template
   B. Knowledge extraction and decision support
2. CERN LHC process
3. Results

Vectorial alarm list representation



$$V^{i,q} = \begin{bmatrix} \boldsymbol{v}_1^{\boldsymbol{i},\boldsymbol{q}} & \boldsymbol{v}_2^{\boldsymbol{i},\boldsymbol{q}} & \boldsymbol{....} & \boldsymbol{v}_{\boldsymbol{N}}^{\boldsymbol{i},\boldsymbol{q}} \end{bmatrix}^{\boldsymbol{T}}$$

with $\boldsymbol{i} \in \boldsymbol{[}1\boldsymbol{,...,}\ \boldsymbol{f}\ \boldsymbol{]}$ and $\boldsymbol{q} \in \boldsymbol{[}1\boldsymbol{,...,k}\ \boldsymbol{]}$

$V^{i,q}$ is the qth alarm list of fault *i*

Weight w^i_j : relevance of alarm *j* to fault *i* compared to faults contained in $S^{\bar{i}}$

$$w^i_j = (2*\alpha_j^{\,i} - 1)*(1-\beta_j^{\,\bar{i}})$$

p^i_j* :** value of alarm ***j for the majority of the k examples of fault ***i***

$$\begin{cases} p^i_j = 1 \text{ if } \dfrac{\sum_{q=1}^{k} v_j^{i,q}}{k} \geq 0.5 \\ p^i_j = 0 \text{ if } \dfrac{\sum_{q=1}^{k} v_j^{i,q}}{k} < 0.5 \end{cases}$$

$$\delta(\mathrm{x}) = \begin{cases} 1 \text{ if } x = 0 \\ 0 \text{ if } x \neq 0 \end{cases}$$

$$\alpha_j^{\,i} = \frac{\sum_{q=1}^{k} \delta(v_j^{i,q} - p^i_j)}{k}$$

Probability that alarm **j** is equal to ***p^i_j*** in the k examples of fault ***i***

$$\beta_j^{\,\bar{i}} = \frac{\sum_{h \in S^{\bar{i}}} \dfrac{\sum_{q=1}^{k} \delta(v_j^{h,q} - p^i_j)}{k}}{m}$$

Probability that alarm **j** is equal to ***p^i_j*** in the examples of faults contained in ***$S^{\bar{i}}$***

With ***$S^{\bar{i}}$*** : a set of faults excluding fault **i** ; m=card($S^{\bar{i}}$)

w^i_j depends on the faults contained in *$S^{\bar{i}}$*

Example : alarm 1

F1	F1	F1	F1	F1	F2	F2	F2	F2	F2	F3	F3	F3	F3	F3
1	0	1	1	0	1	1	1	1	1	0	0	0	0	0

F1 compared with F2 and F3 :

$$p_1^1 = 1 \quad \alpha_1^1 = \frac{3}{5} \quad \beta_1^1 = \frac{5}{10} \quad \omega_1^1 = (2 * \frac{3}{5} - 1) * (1 - \frac{5}{10}) = 0.2 * 0.5 = 0.1$$

F1 compared with F2 :

$$p_1^1 = 1 \quad \alpha_1^1 = \frac{3}{5} \quad \beta_1^1 = 1 \quad \omega_1^1 = (2 * \frac{3}{5} - 1) * (1 - 1) = 0$$

F2 compared with F3 :

$$p_1^2 = 1 \quad \alpha_1^2 = 1 \quad \beta_1^2 = 0 \quad \omega_1^2 = (2 * 1 - 1) * (1 - 0) = 1$$

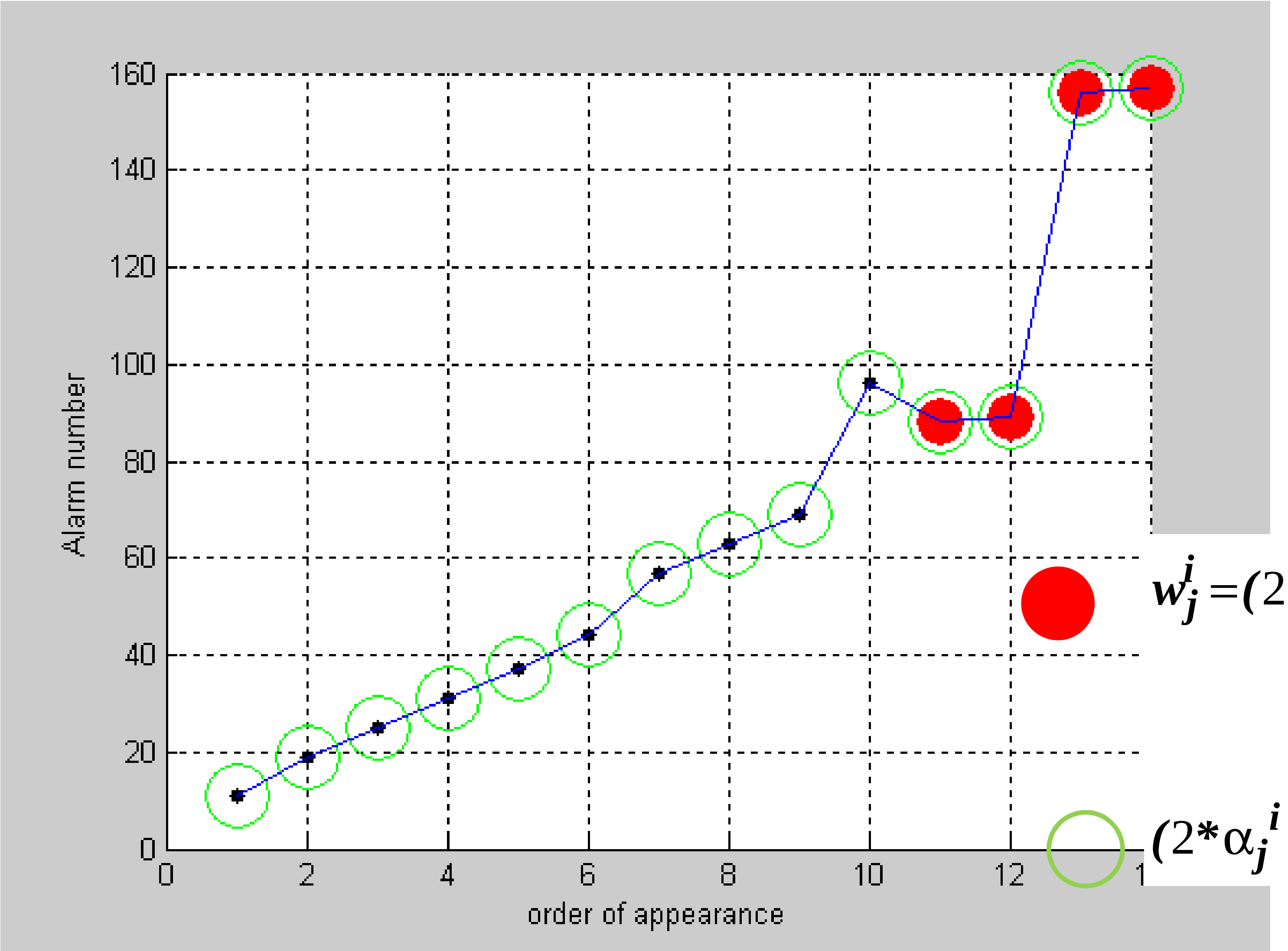
Outline

1. Method
   A. Template extraction
      - sequential template
      - vectorial template
      - weighted sequential template
   B. Knowledge extraction and decision support

2. CERN LHC process

3. Results

Weighted sequential template : T^i



$w_j^i = (2*\alpha_j^i - 1)*(1-\beta_j^{\bar{i}})$

$(2*\alpha_j^i - 1) = 1$

Outline

1. Method
   A. Template extraction
      - sequential template
      - vectorial template
      - weighted sequential template
   B. Knowledge extraction and decision support

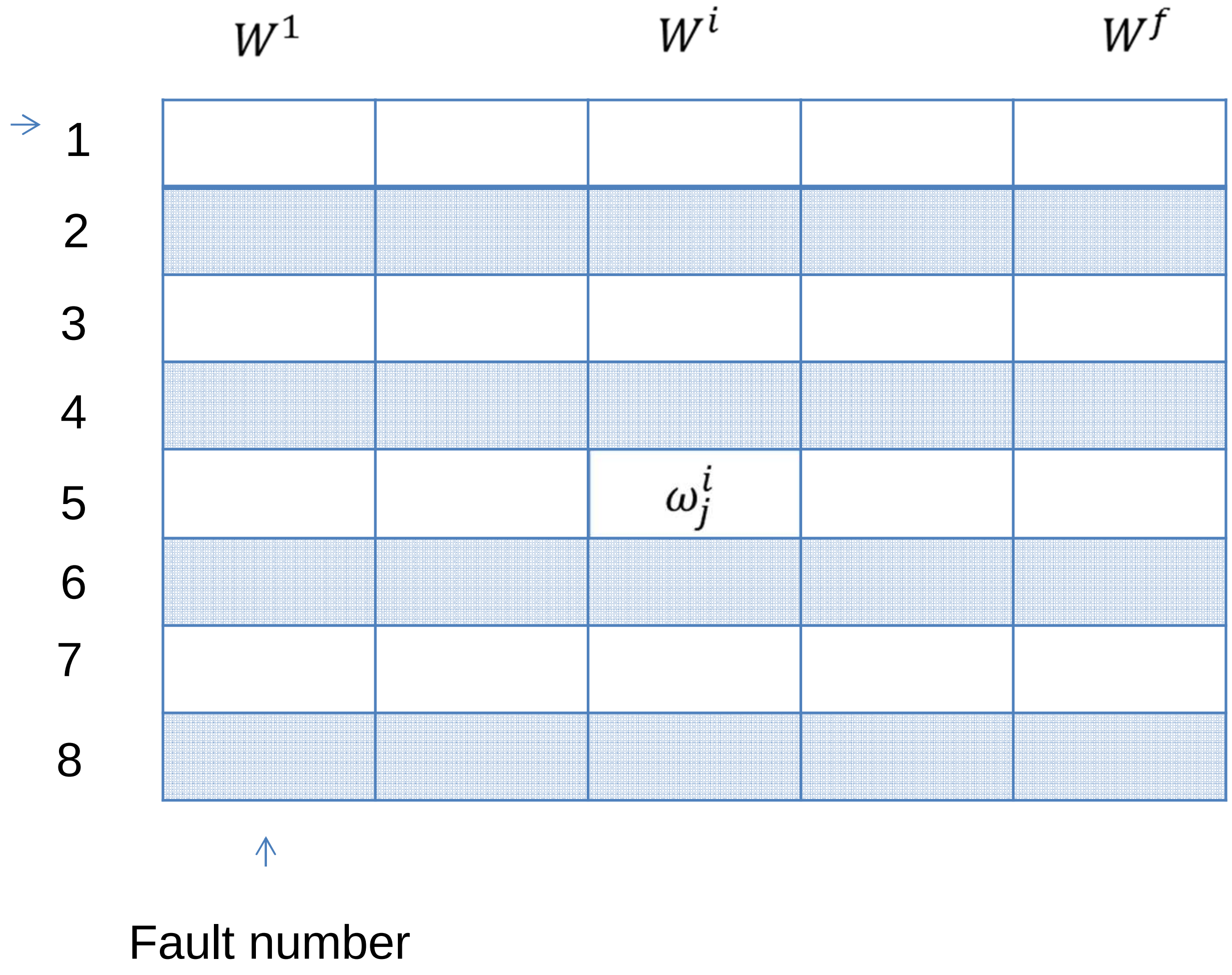
2. CERN LHC process

3. Results

Weigth matrix representation

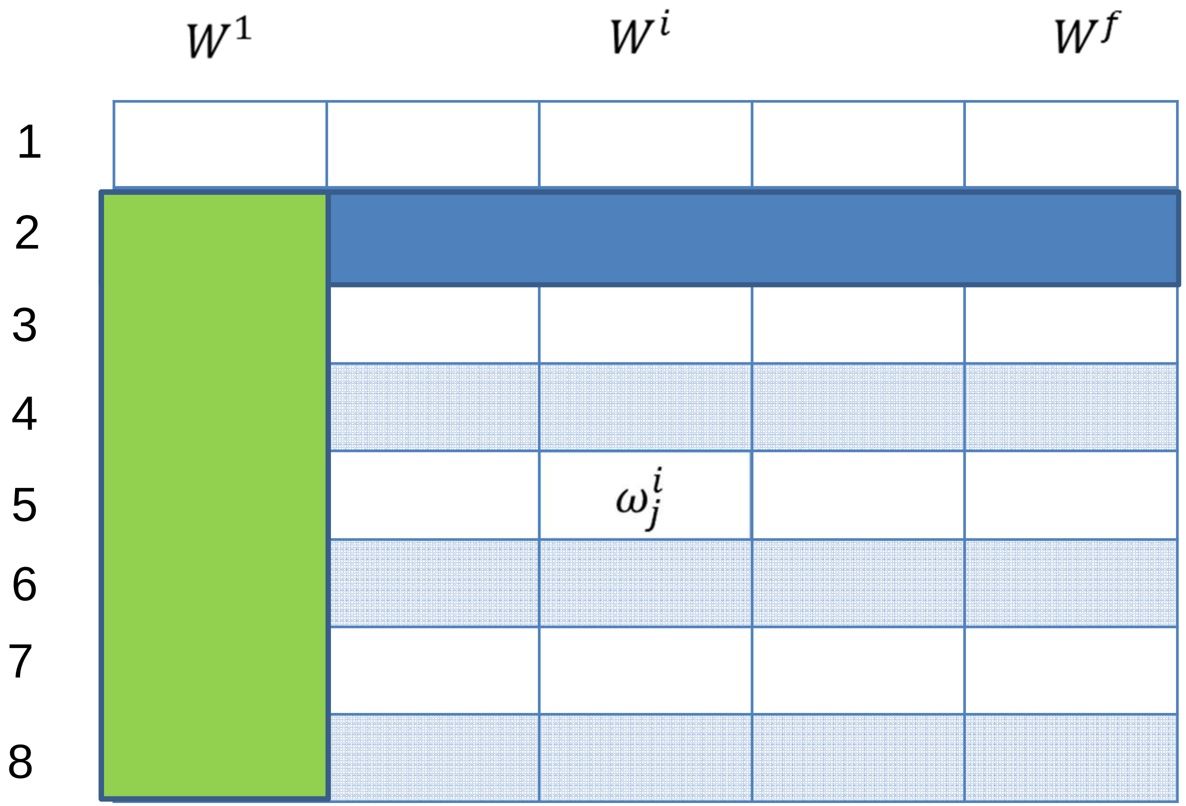
$$W^i = \begin{bmatrix} w_1^i \\ w_2^i \\ \\ w_N^i \end{bmatrix}$$

Alarm number



Weigth matrix representation

W^i with ***S^i = {1,…,f}\i***



$$\max(w_j^{\,i}) \text{ for } i = 1 \text{ to } f$$

$$\max(w_j^{\,i}) \text{ for } j = 1 \text{ to } N$$

Knowledge extraction : fault diagnosability

Definitions :

- An alarm ***j*** is specific to a fault ***i*** compared to a subset of faults $\boldsymbol{S^i}$ if $w_j^i = 1$

- if a fault ***i*** has at least one specific alarm, it can be distinguished from the faults contained in $\boldsymbol{S^i}$, from its vectorial representation

Discriminability between pairs of faults (i,k)

$$W^i = \begin{bmatrix} w_1^i \\ w_2^i \\ \vdots \\ w_N^i \end{bmatrix}$$

With $\boldsymbol{S^i = \{k\}}$

$$W^i = \begin{array}{c} \\ \\ \\ j \\ \\ \\ \\ \end{array} \begin{array}{|c|} \hline \\ \hline \\ \hline \\ \hline \mathbf{1} \\ \hline \\ \hline \\ \hline \\ \end{array}$$

Alarm ***j*** is specific to fault ***i*** compared to fault ***k***

Alarm ***j*** can distinguish fault ***i*** from fault ***k***

For k=1 to f except i,

Calculate W^i using $S^i = \{k\}$

If $\max_{j=1:N}(w_j^i) = 1$ fault i can be distinguished from fault k

else it can be confused with fault k : $k \in \overline{F}_i$

End

$\overline{F}_i = \{\ \}$ **Fault i is diagnosable from its alarm list**

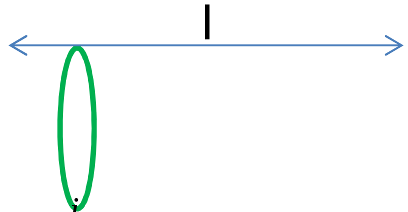
Decision support in operation : similarity between unknown alarm sequence S and template T^i

Unweighted similarity between S and template T^i (length q)

$$sim_{NWA}(S, T^i) = \frac{\sum_{p=1}^{l} \sigma(t_p^{a^i}, s_p^a)}{l}$$

=0

$$C = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$



$\sigma(t_p^{a^i}, s_p^a)$

Weighted similarity between S and template T^i (length : q)

$$WS(S, T^i) = \frac{\sum_{p=1}^{l} \sigma(t_p^{a^i}, s_p^a)}{\sum_{h=1}^{q} C(t_h^i, t_h^i) + (l - q)}$$

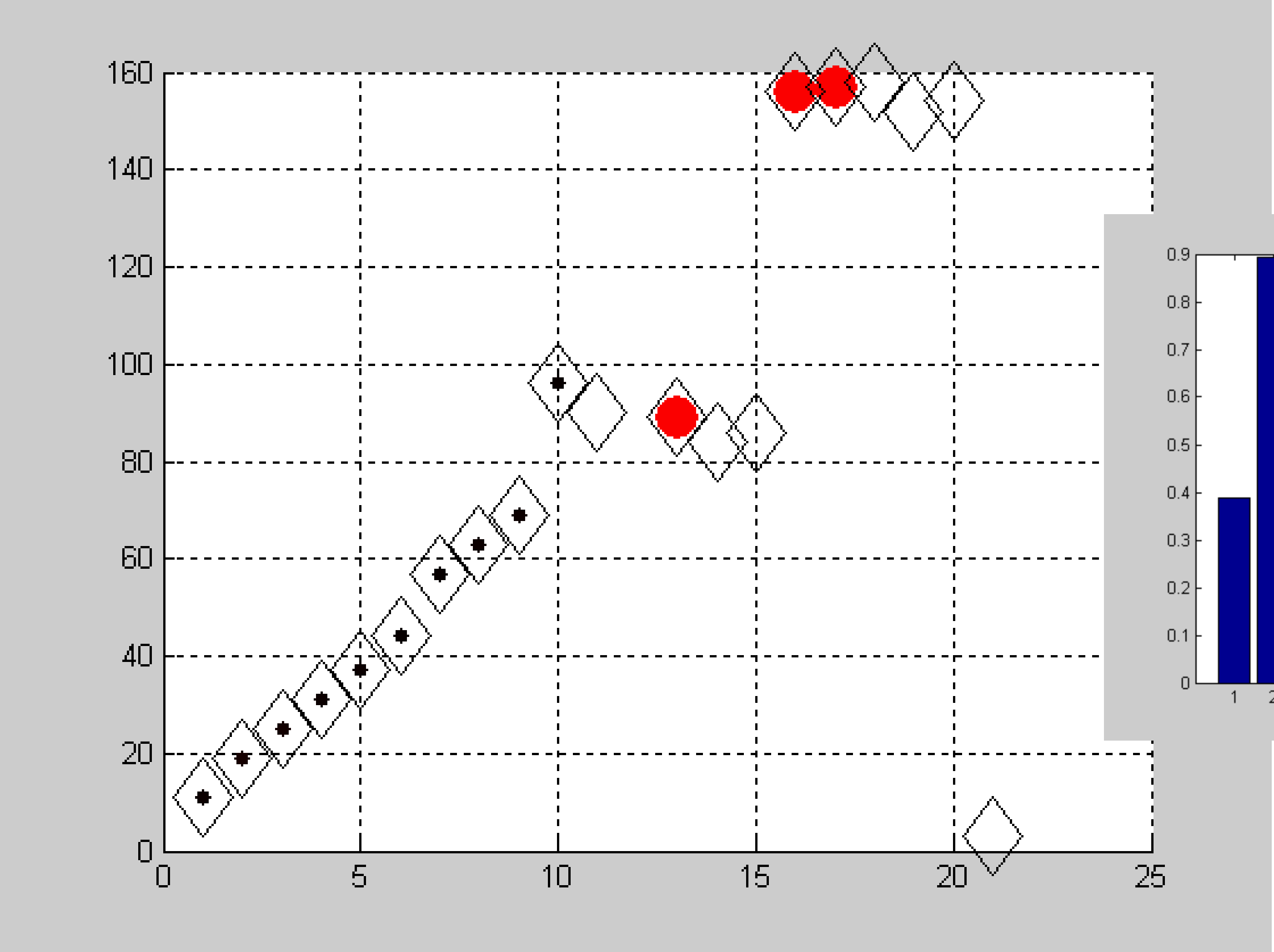
$$C = \begin{bmatrix} 1+\alpha w_1^i & -1 & -1 \\ -1 & 1+\alpha w_j^i & -1 \\ -1 & -1 & 1+\alpha w_N^i \end{bmatrix}$$

=0

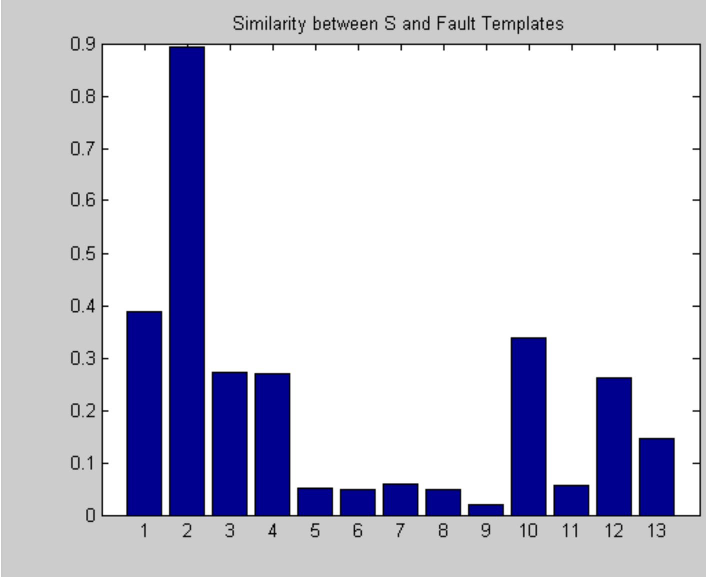
Alarms present in T^i

number of gaps added

Decision support in operation : information display



similarity between S and fault templates



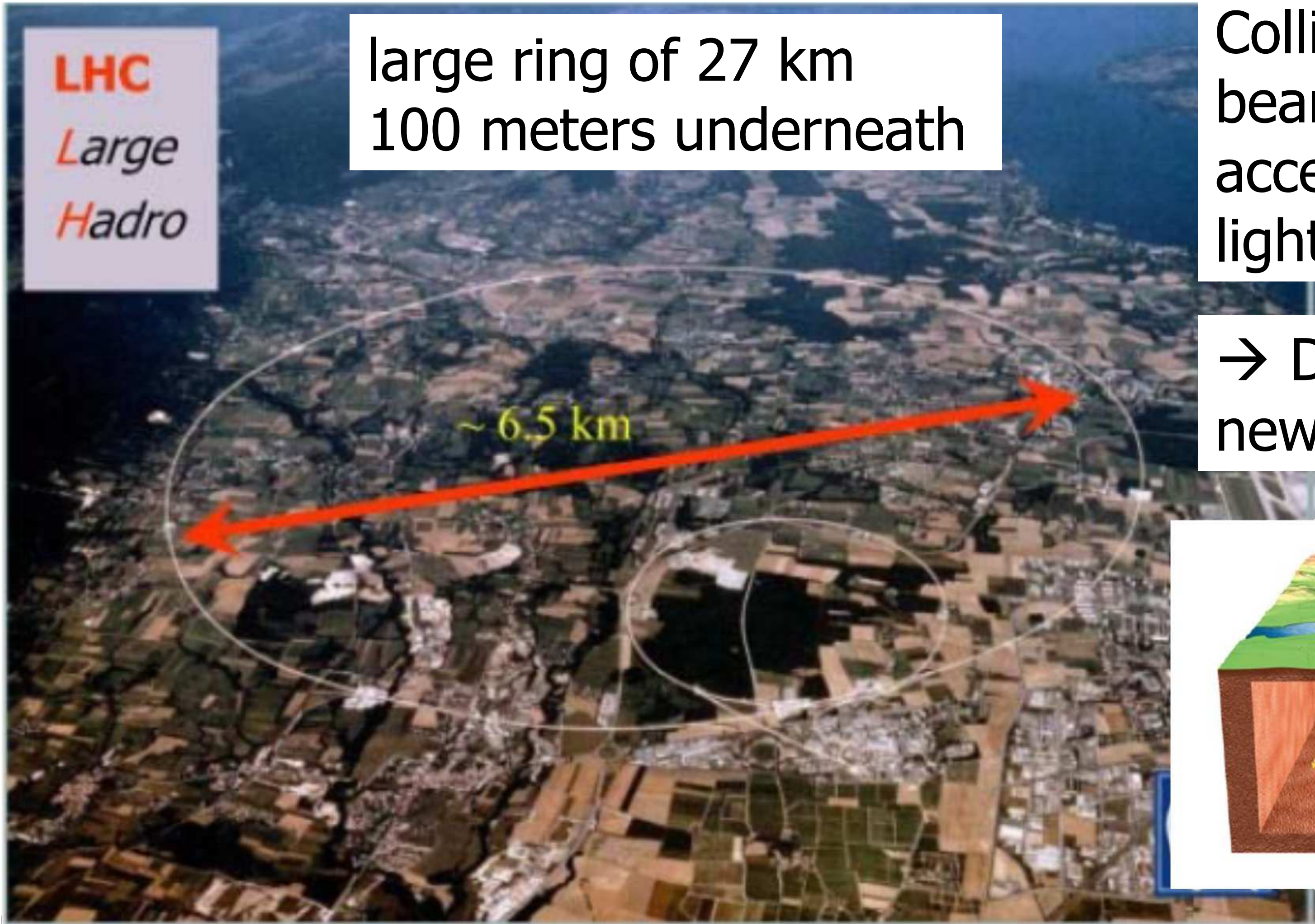
Outline

1. Method
   A. Template extraction
      - sequential template
      - vectorial template
      - weighted sequential template
   B. Knowledge extraction and decision support

2. CERN LHC process

3. Results

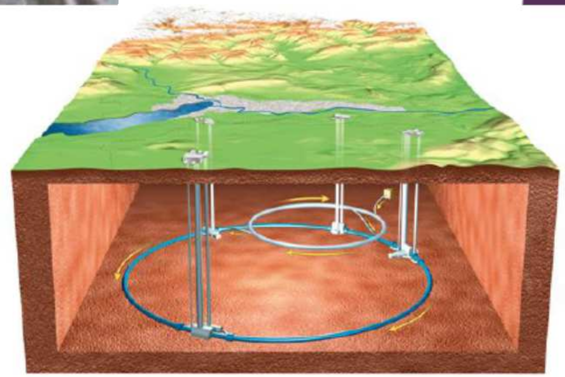
Large Hadron Collider in CERN



large ring of 27 km
100 meters underneath

Collision of 2 beams of protons accelerated at the light speed

→ Discovery of new particules



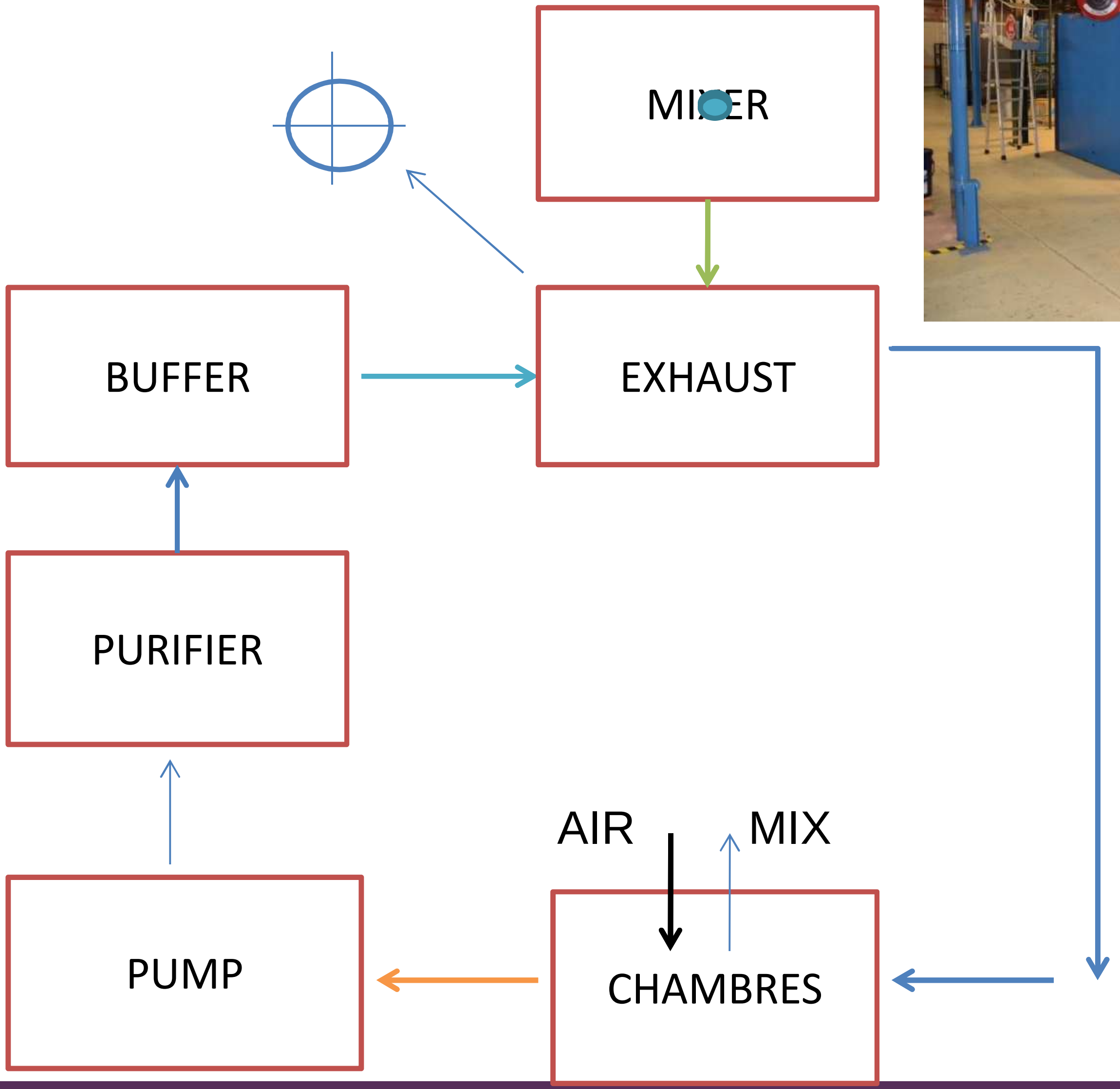
Large Hadron Collider systems

LHC :

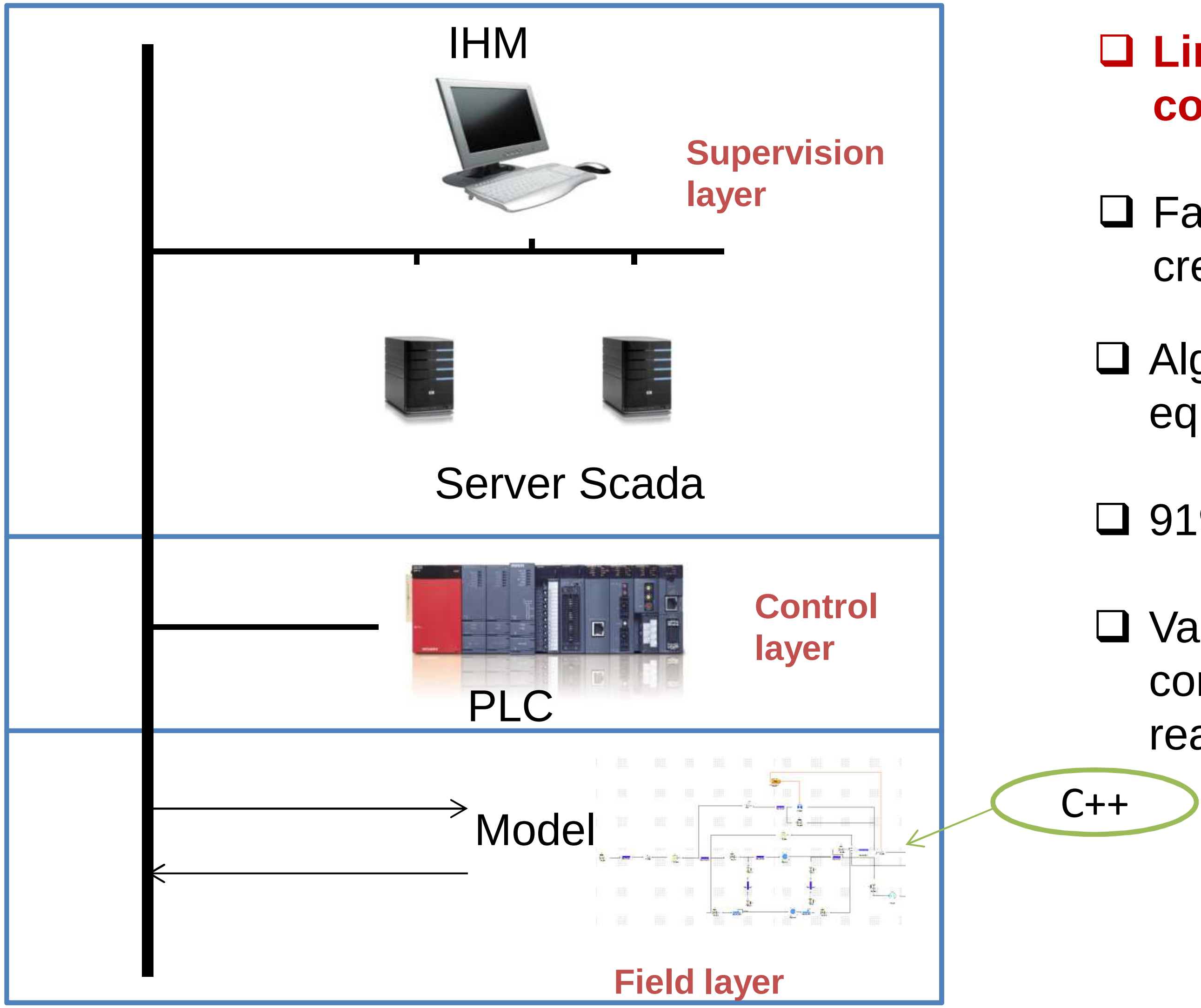
- the biggest particle accelerator in the world
- formed of many different systems
- all controlled using the same SCADA architecture

The gas system : supply the detection chambers with a precise mixture of pure gases

CERN LHC process : the gaz system

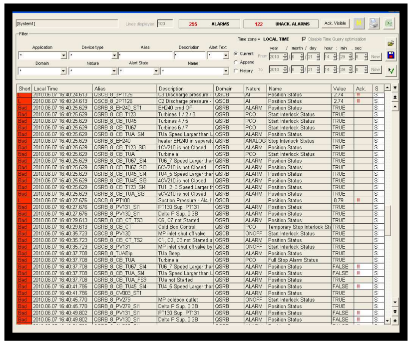


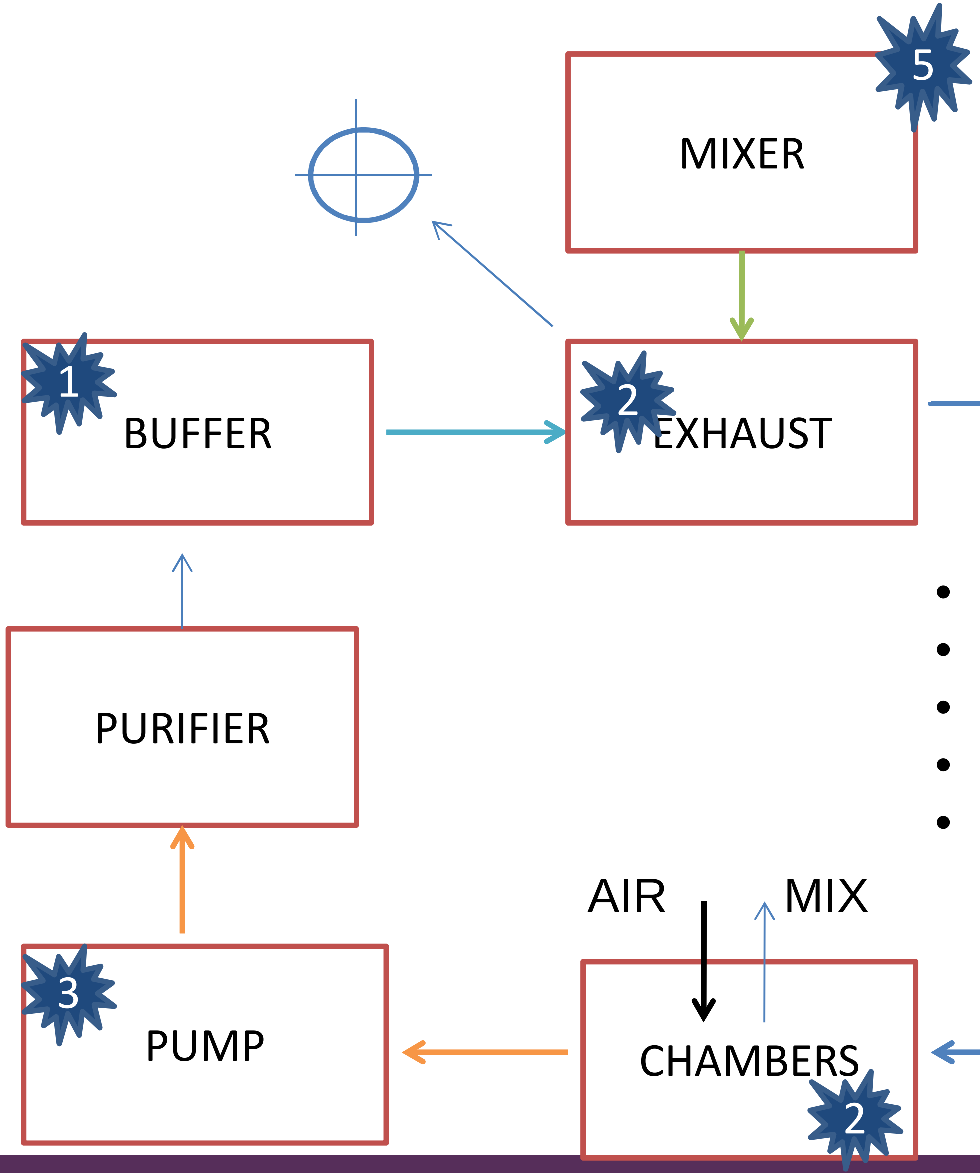
Accurate simulator of the gas system



- **Linked to the real control system**
- Faults can be created
- Algebro-differential equations
- 9195 equations
- Validated from comparison with the real system

f= 13 faults generated





k=6 alarm lists per fault
N=160 alarms
(triggered at least once)

- Leaks
- blockage of controllers
- stoppage of pumps
- broken bubblers (safety device)
- sensor problems leading to regulation issues

65/78 Alarm Floods
(>10 alarms in 10 min)

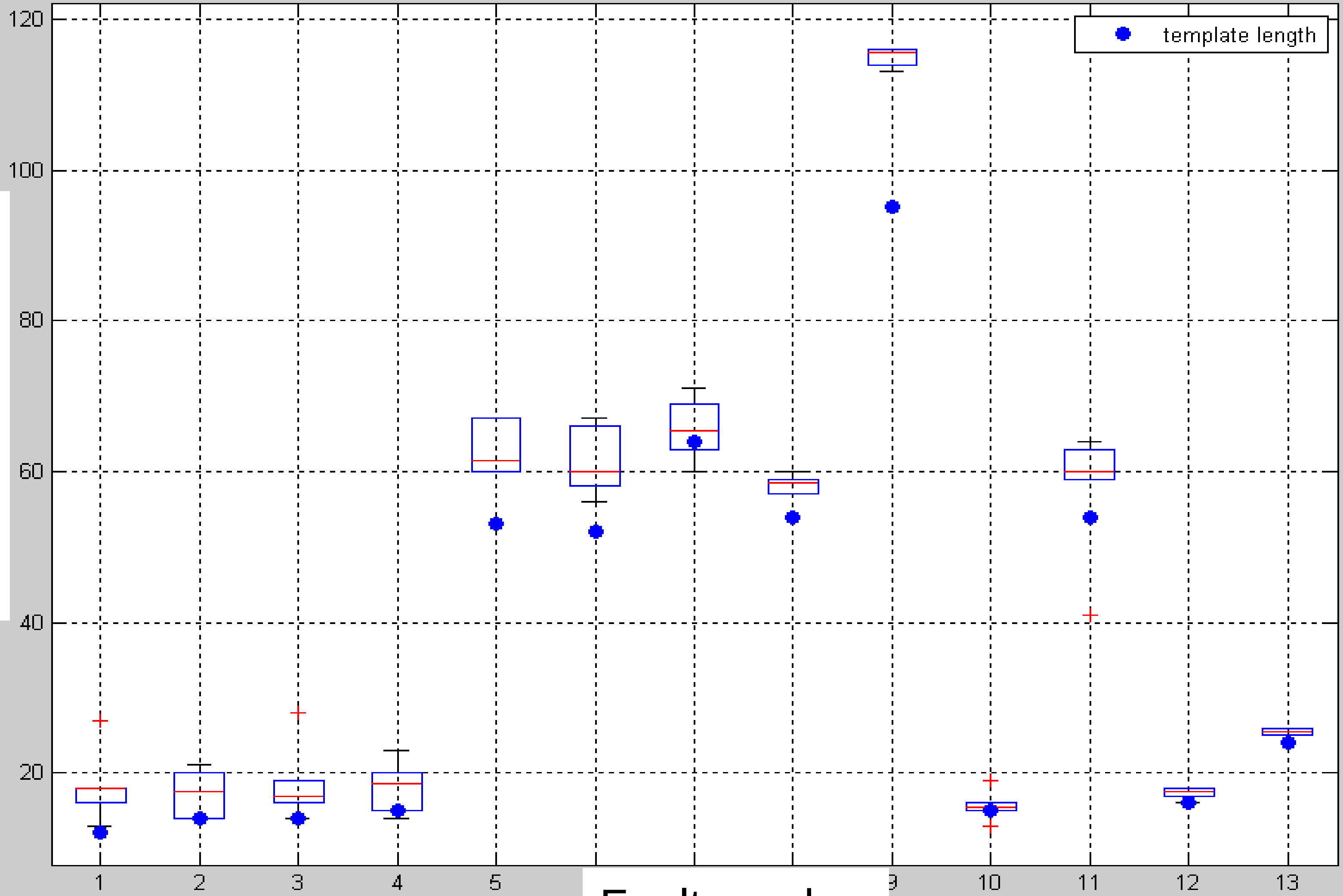
Outline

1. Method
   A. Template extraction
      - sequential template
      - vectorial template
      - weighted sequential template
   B. Knowledge extraction and decision support

2. CERN LHC process

3. Results

Alarm lists length – Template length



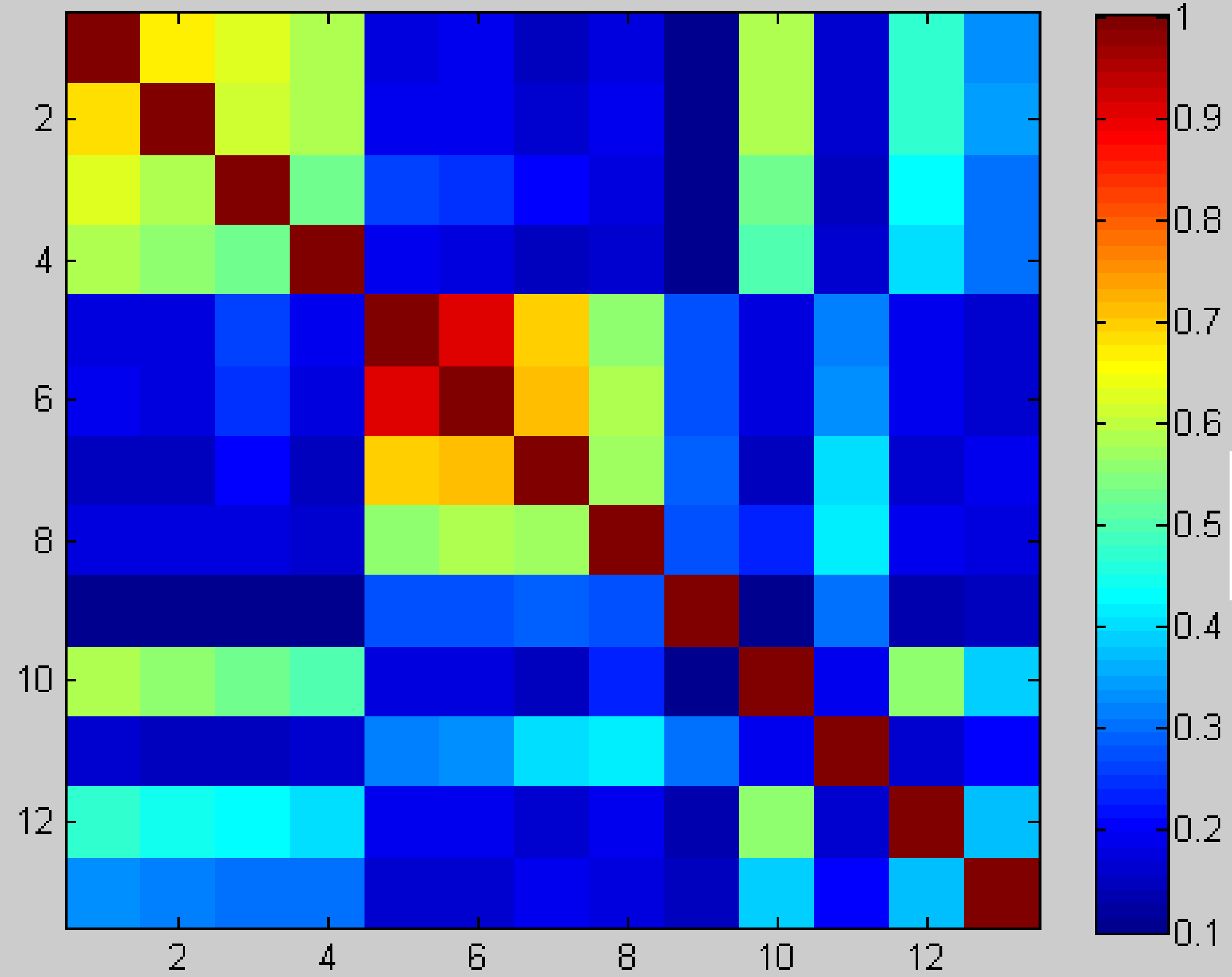
Number of alarmrs

Fault number

0.5	0.8	0.6	0.8	1	1	1	1	1	0.5	1	0.6	0.8

Percentage of alarm floods

Similarity between sequential fault template : unweighted similarity

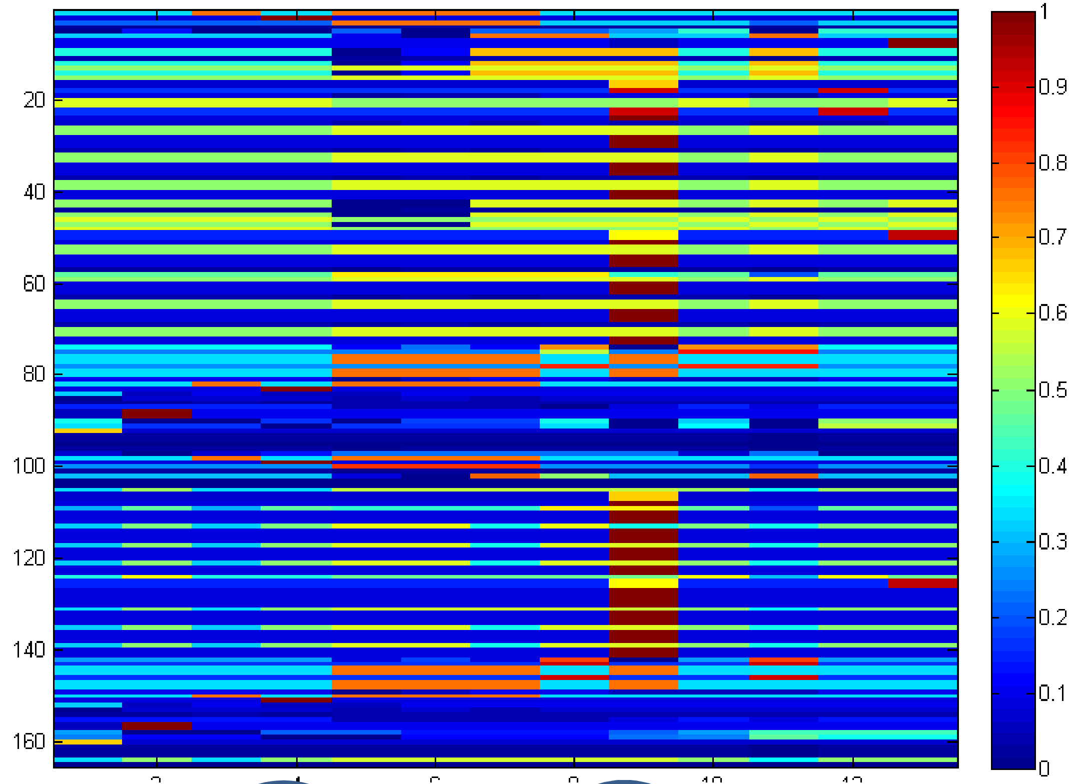


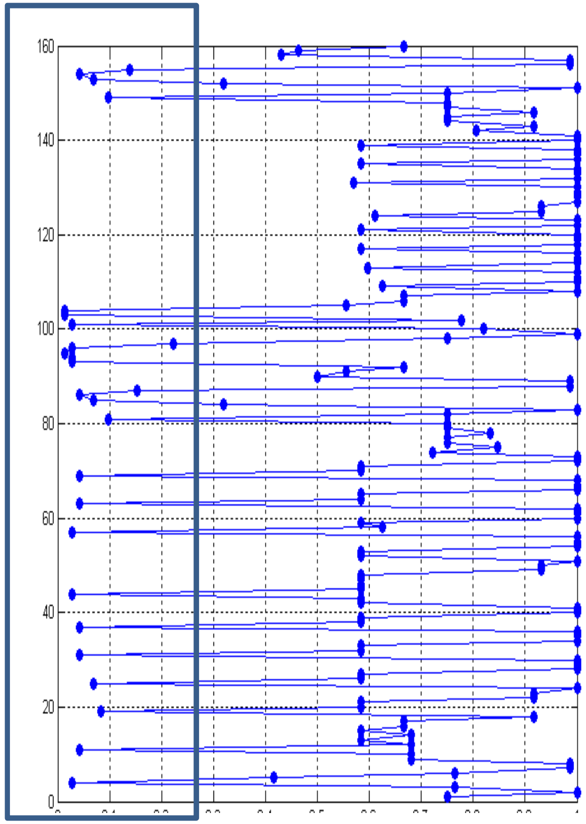
19 18 19 19 63 62 66 59 115 16 58 18 26

Alarm list averaged length

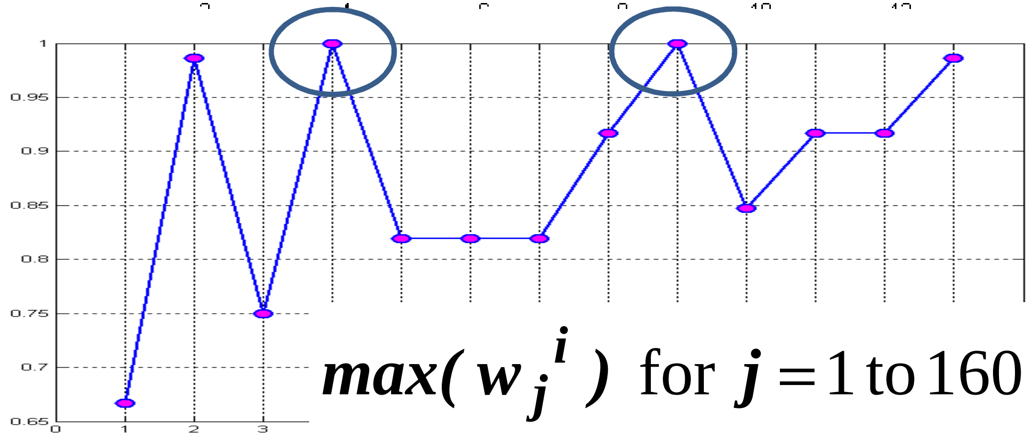
$$C = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$

Weigth analysis $w_j^{\,i}$ with $S^{\mathrm{i}} = 13\,\text{faults except}\; i$





$\max(w_j^{\,i})$ for $i = 1$ to 13

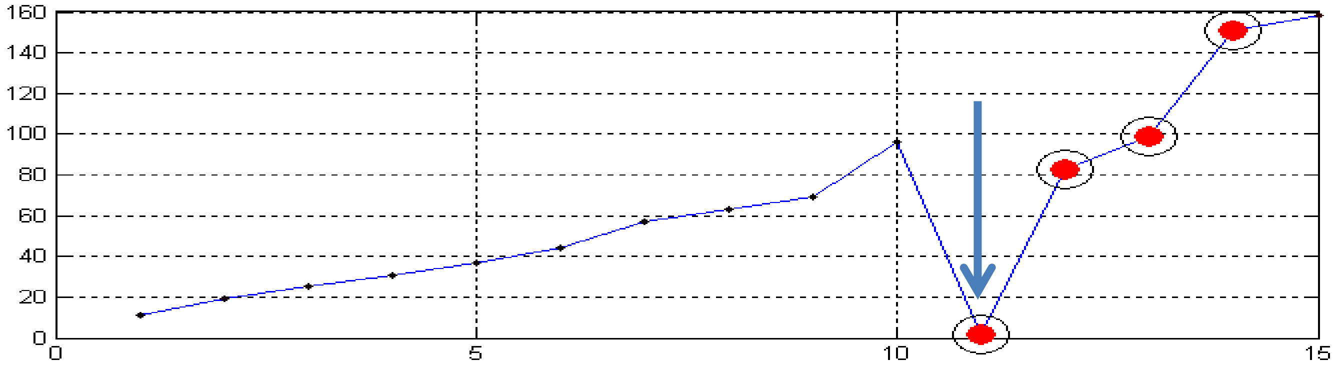


$\boldsymbol{max(\,w_j^{\,i}\,)}$ for $\boldsymbol{j} = 1$ to 160

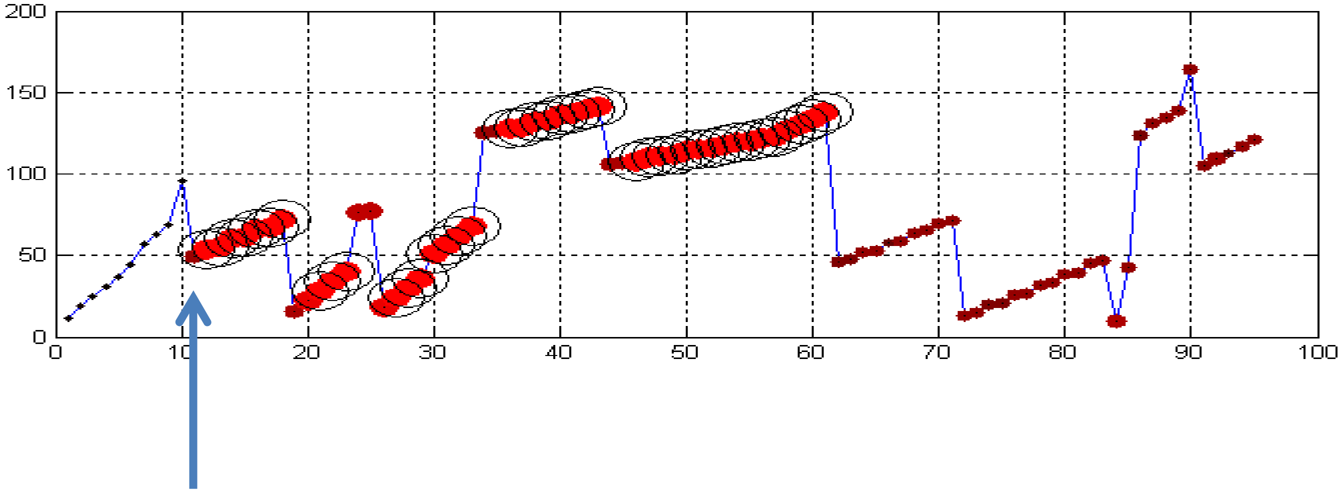
- 17% of alarms are irrelevant
- Faults 4 and 9 can be diagnosed

Weighted sequential fault template

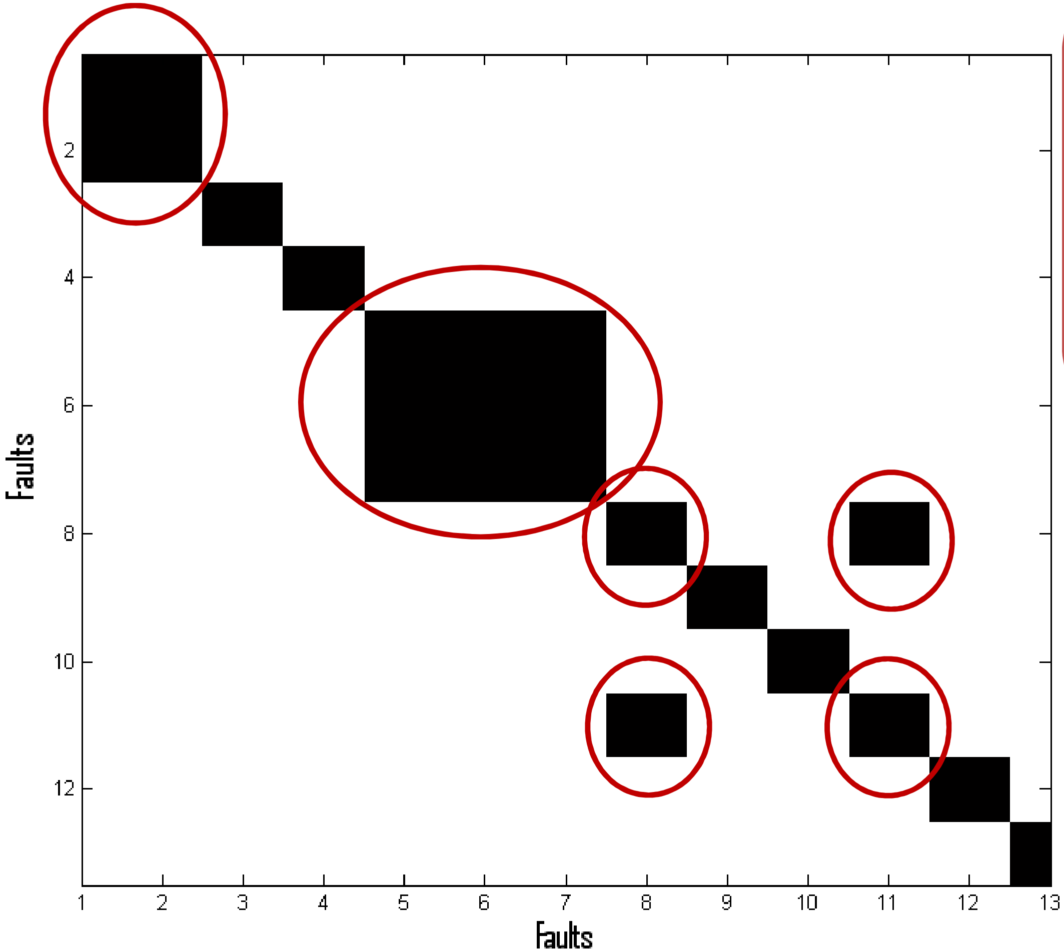
Fault 4



Fault 9



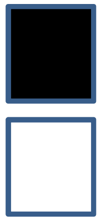
Fault diagnosability (vectorial template)



If $\max_{j=1:N}(w_j^i) = 1$

where $S^{\bar{i}} = \{k\}$

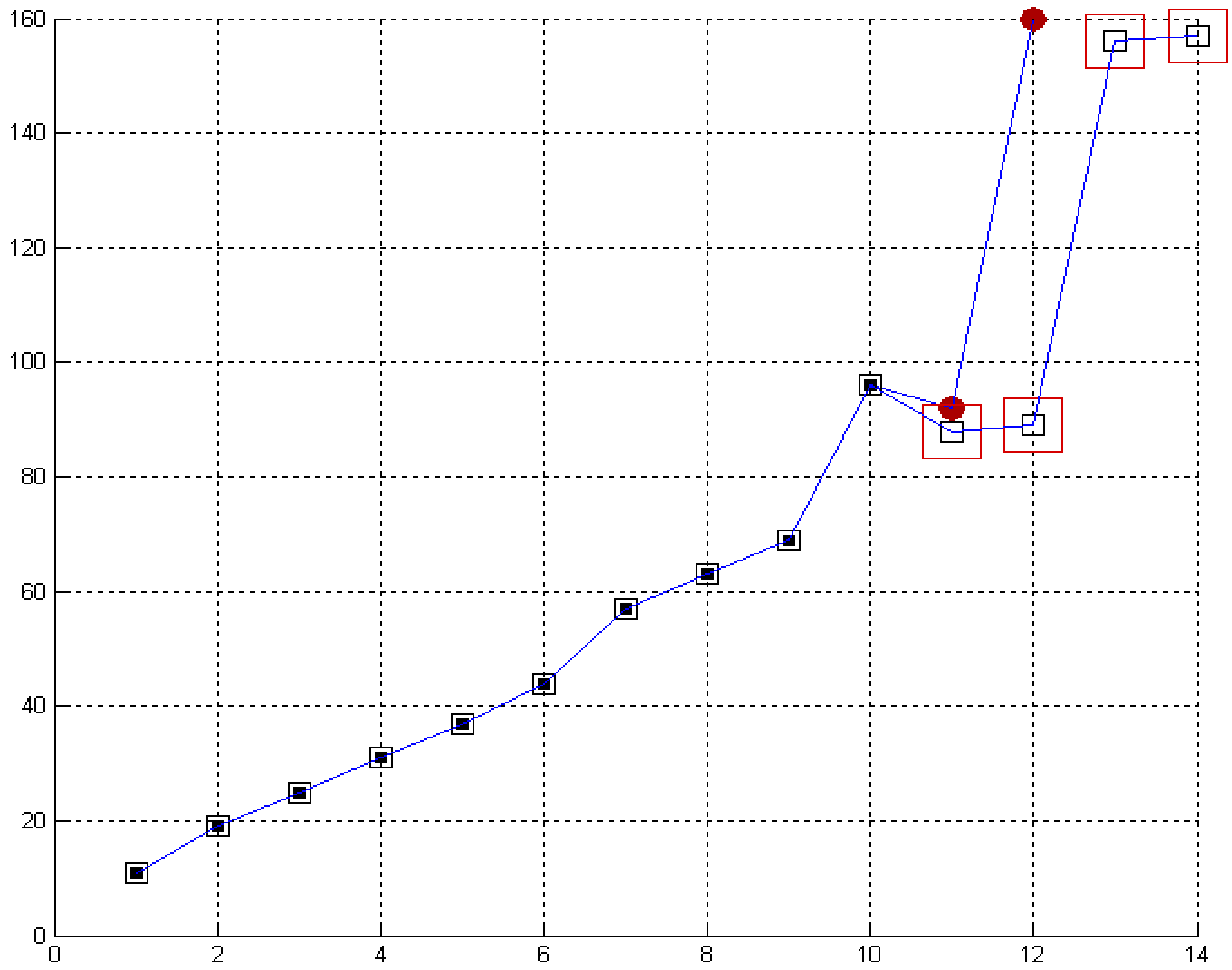
→ fault i can be distinguished from fault k



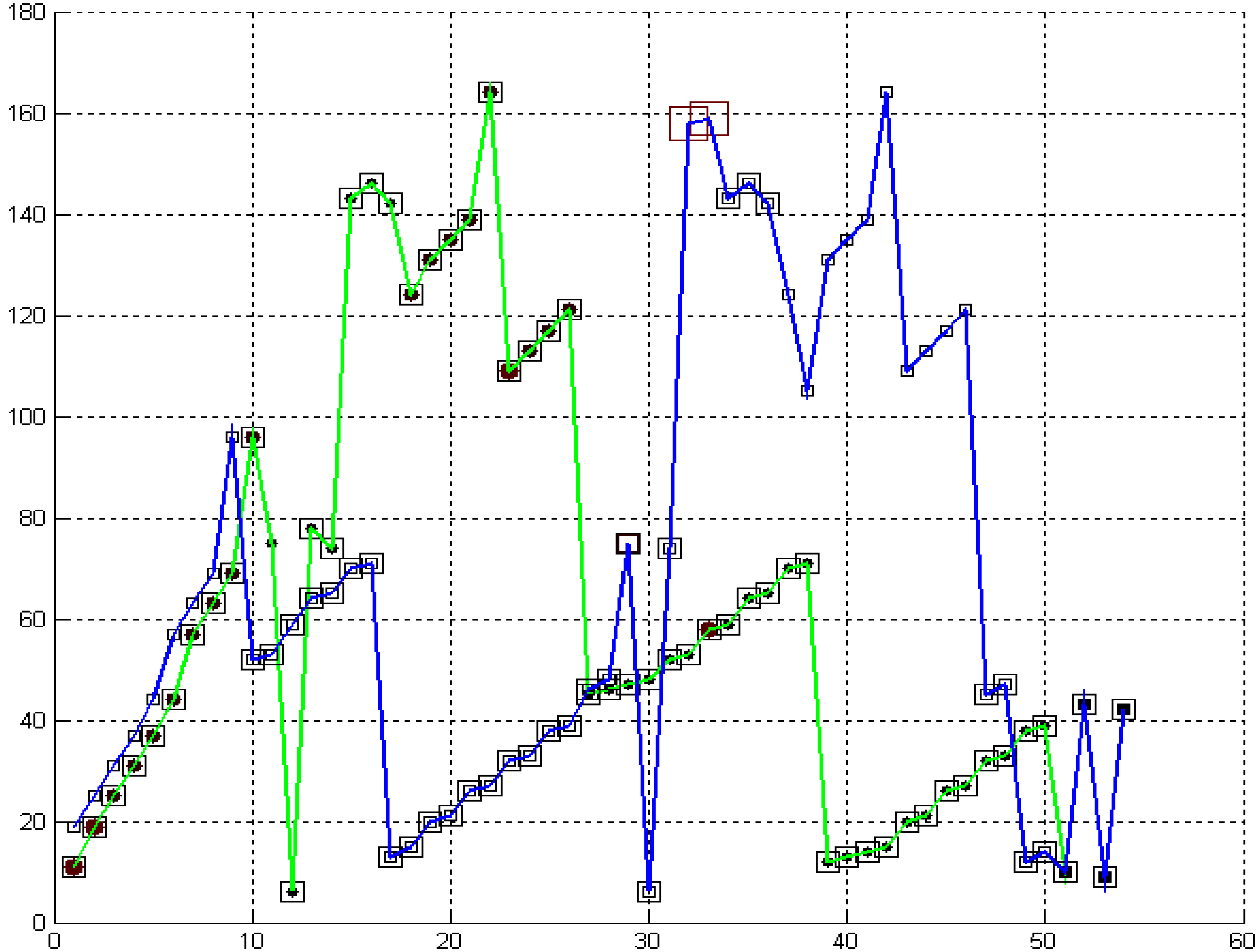
Faults **3, 4, 9, 10, 12, 13** are **diagnosable from their alarm vectors**
Confusion : {1,2}, {5,6,7} , {8,11}

Fault 1/ Fault 2

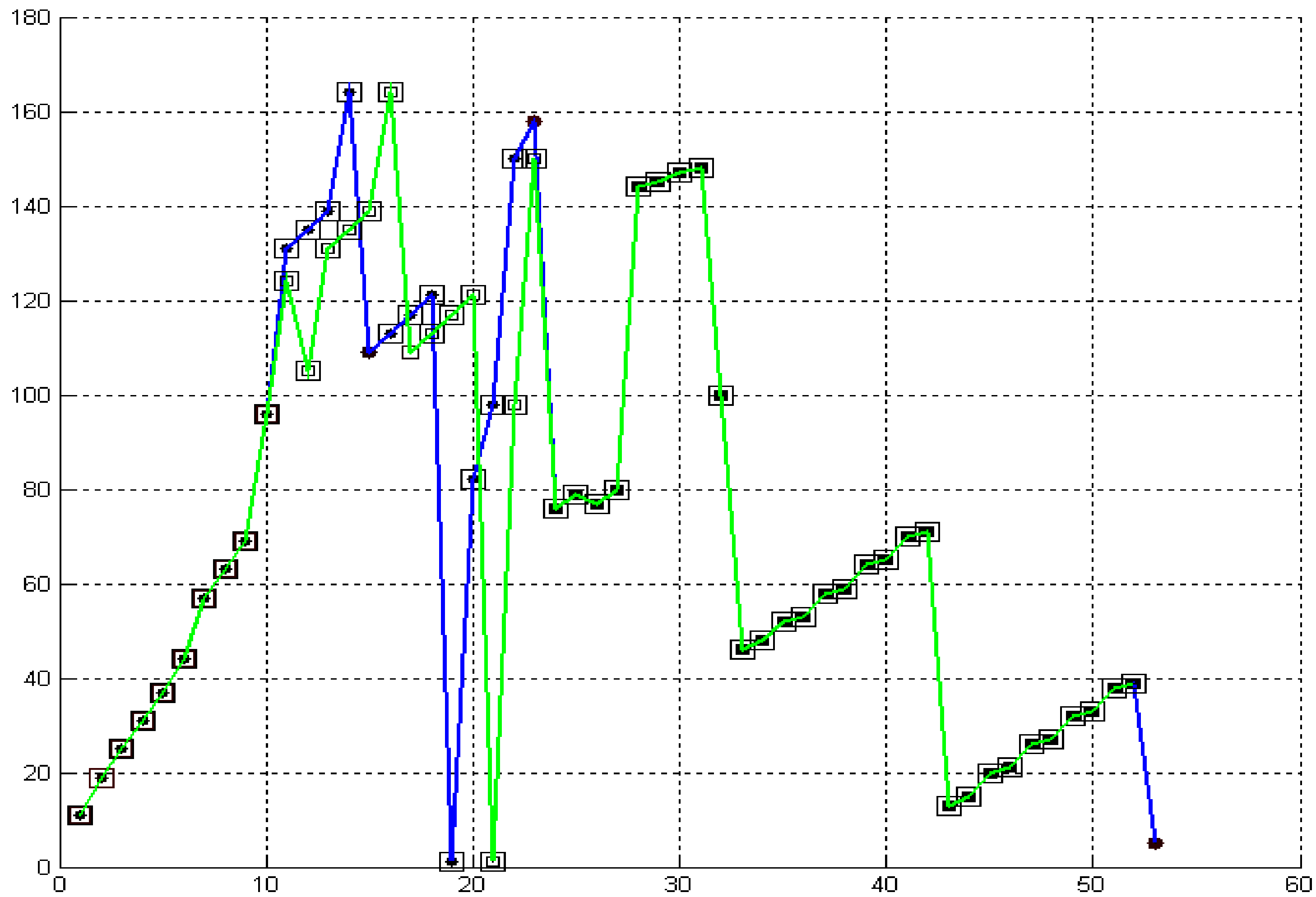
$$w_j^i \text{ with } s^{\bar{i}} = \{k\}$$



Fault 8/ Fault 11



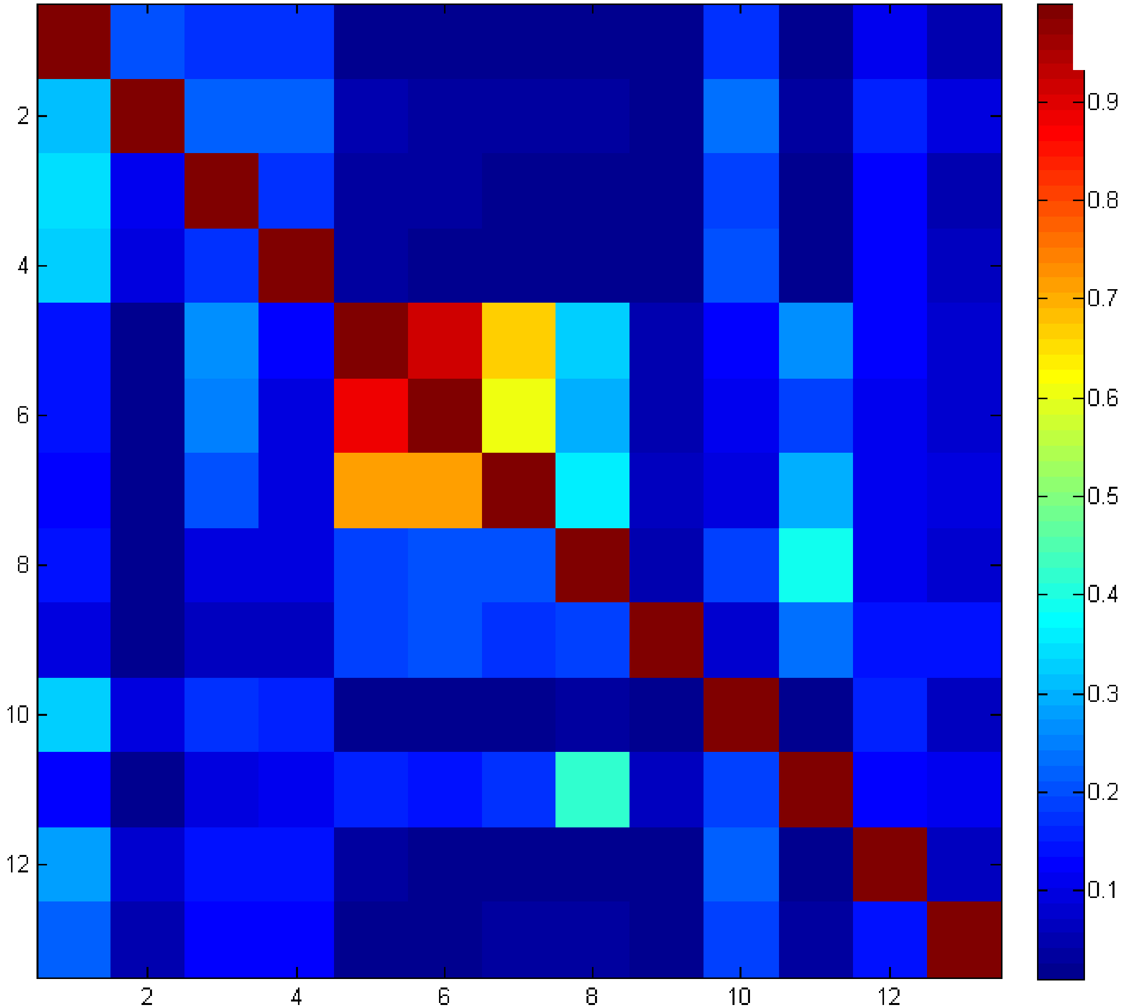
Fault 5/ Fault 6



Leak in Mixer/ leak in Exhaust→ Lack of instrumentation

Similarity between sequential fault templates : weighted similarity

$$WS(S\ ,S^{i}) = \frac{\sum_{p=1}^{l} \sigma(s_p^{a^i}, s_p^{a})}{\sum_{h=1}^{q} C(s_h^{i}\ , s_h^{i}\ ) + (l-q)}$$

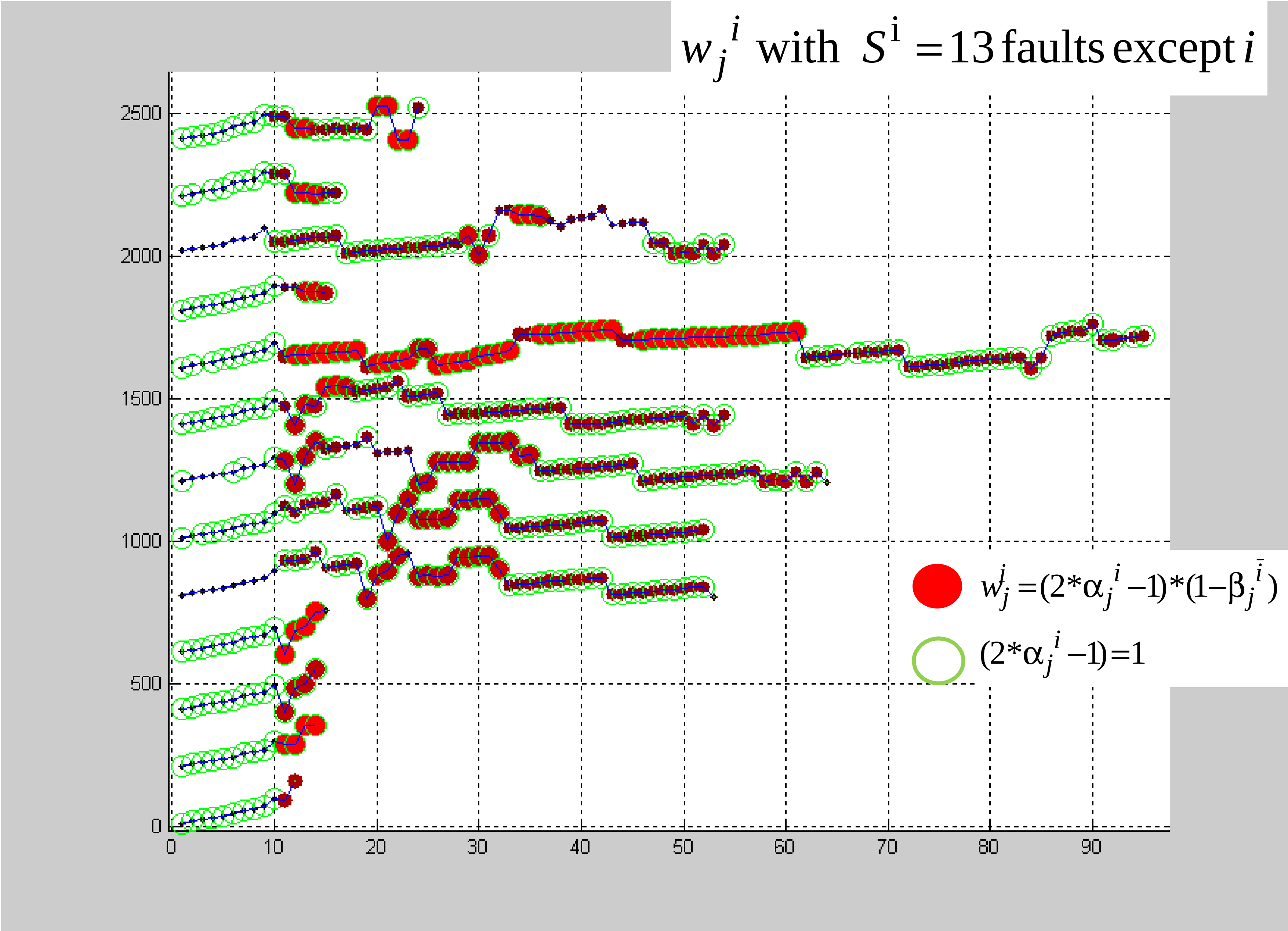


$$C = \begin{bmatrix} 1+\alpha w_1^i & -1 & -1 \\ -1 & 1+\alpha w_j^i & -1 \\ -1 & -1 & 1+\alpha w_N^i \end{bmatrix}$$

α=10

$$\boldsymbol{w_j^i} \text{ with } \boldsymbol{s^{\bar{i}}} = \{\boldsymbol{l}\}$$

Fault templates for i =1 to 13



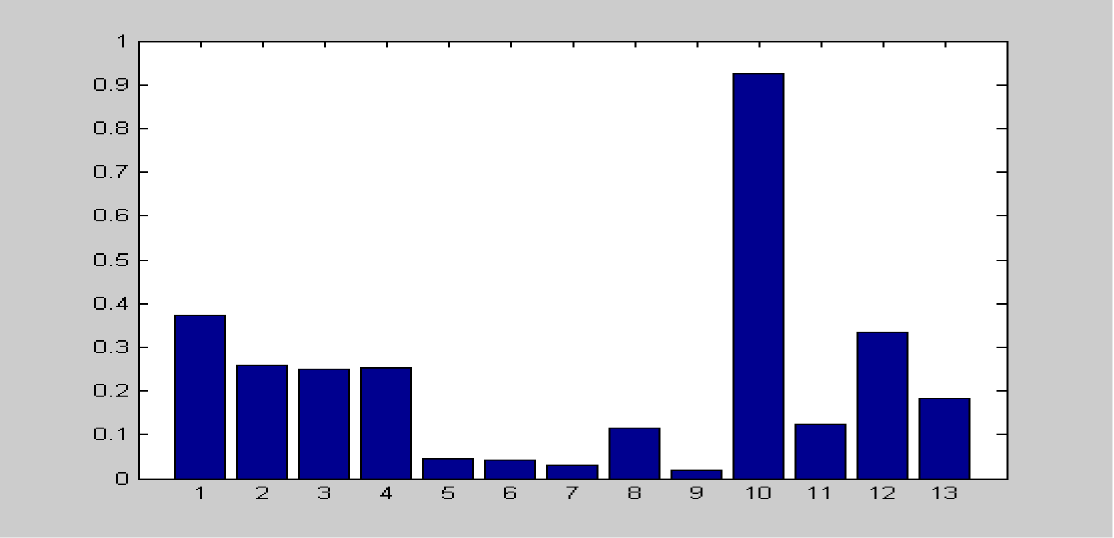
Decision support in operation → diagnose unknown sequence S

For i=1 to 13

$$WS(S\ ,T^i) = \frac{\sum_{p=1}^{l} \sigma(t_p^{a^i}, s_p^a)}{\sum_{h=1}^{q} C(t_h^i, t_h^i) + (l-q)}$$

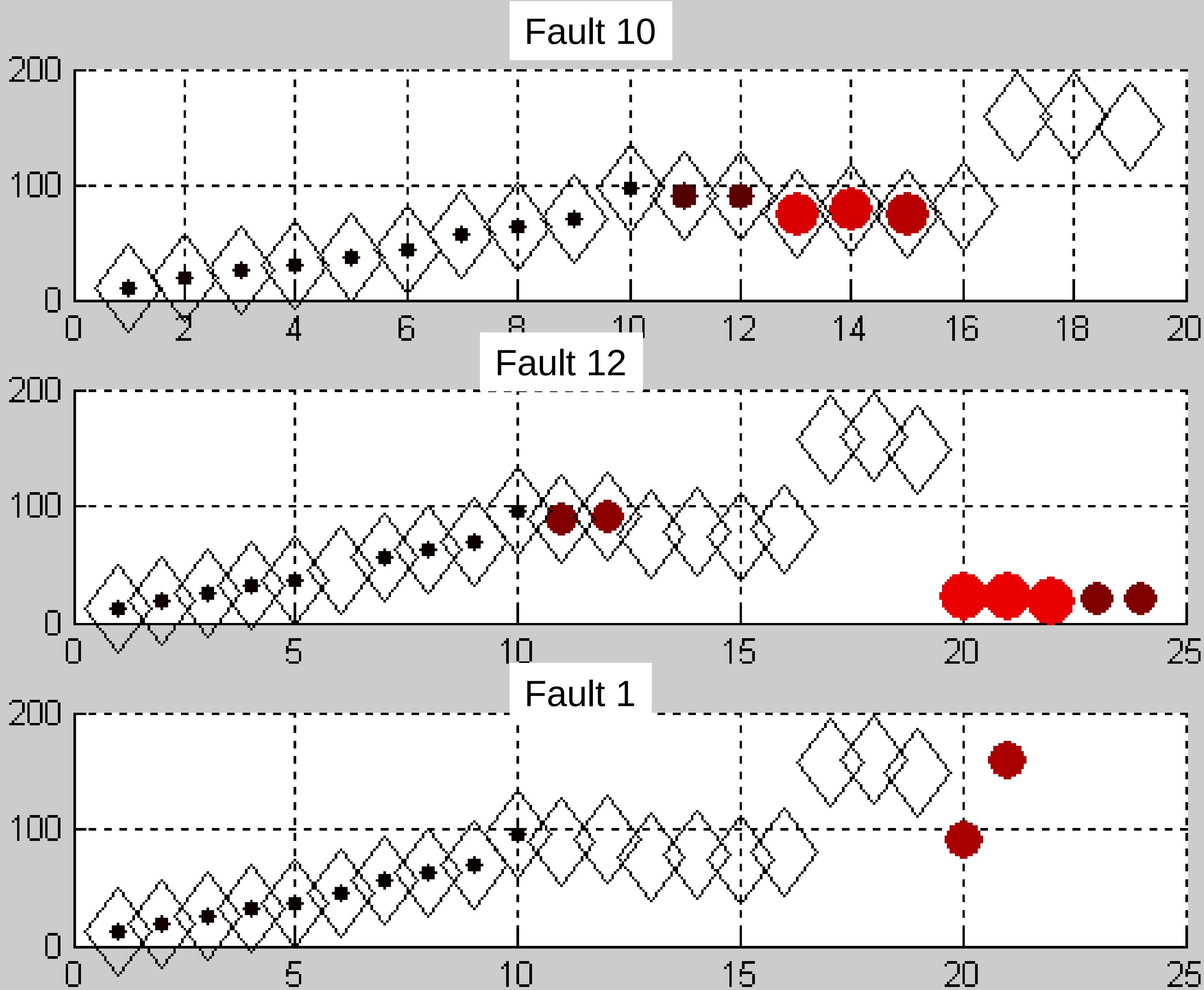
$w_j^{\ i}$ with $S^{\mathrm{i}} = 13$ faults except i

The most similar fault: 10



Weighted similarity between S, and Ti

Fault with a similarity higher than 0.3 :
1, 10, 12



Fault 10
200
100
0
0
2
4
6
8
10
12
14
16
18
20
Fault 12
200
100
0
0
5
10
15
20
25
Fault 1
200
100
0
0
5
10
15
20
25

Classification accuracy using the first nearest neighboor

For each of the 78 sequences

Classification : Assign unknown sequence S to the fault whose template is the most similar (weighted similarity)

Validation : leave one out

- remove S from the data set
- extract fault templates Ti without S

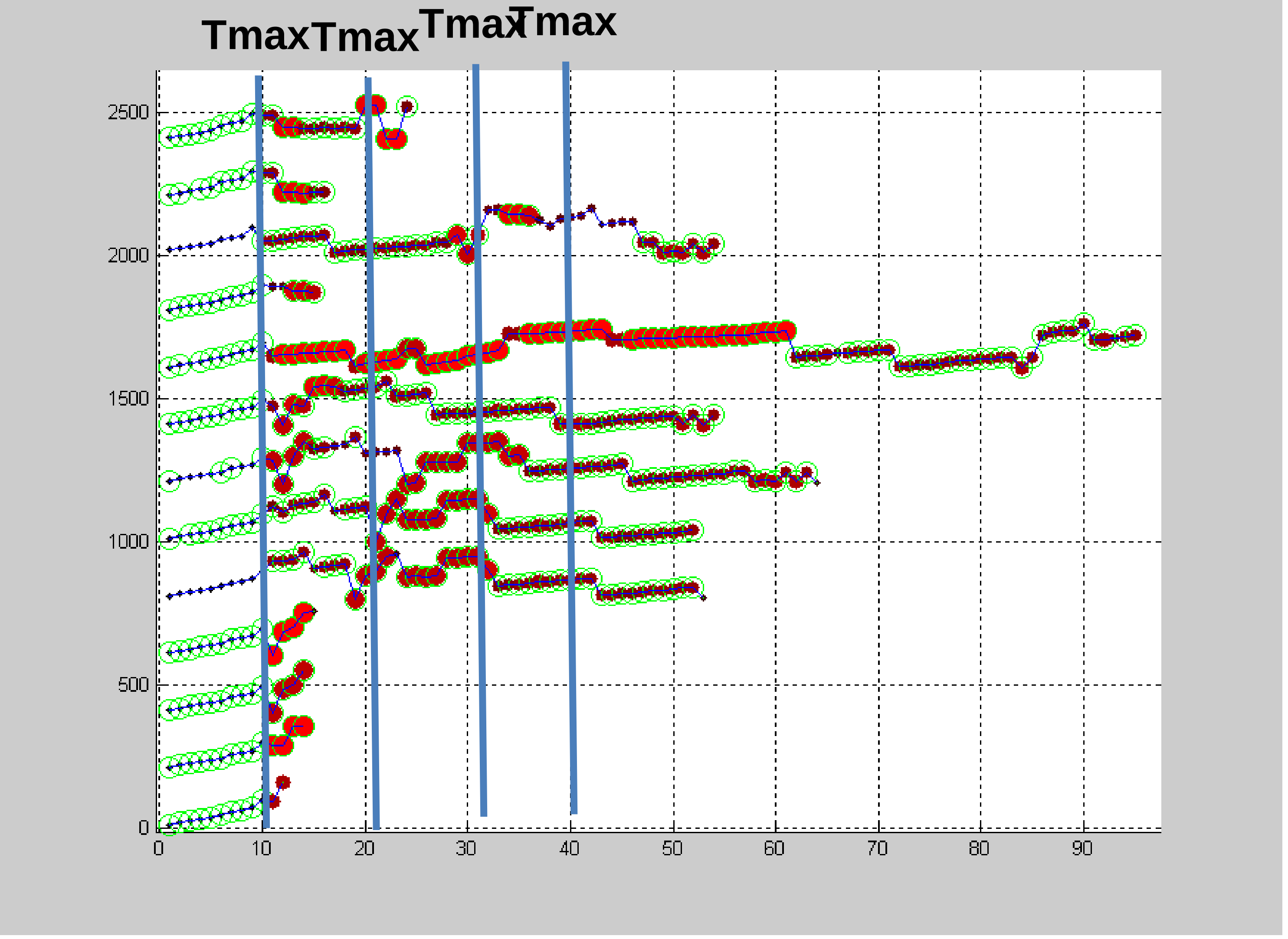
Results :

Vectorial representation ; Hamming distance : **54/78**

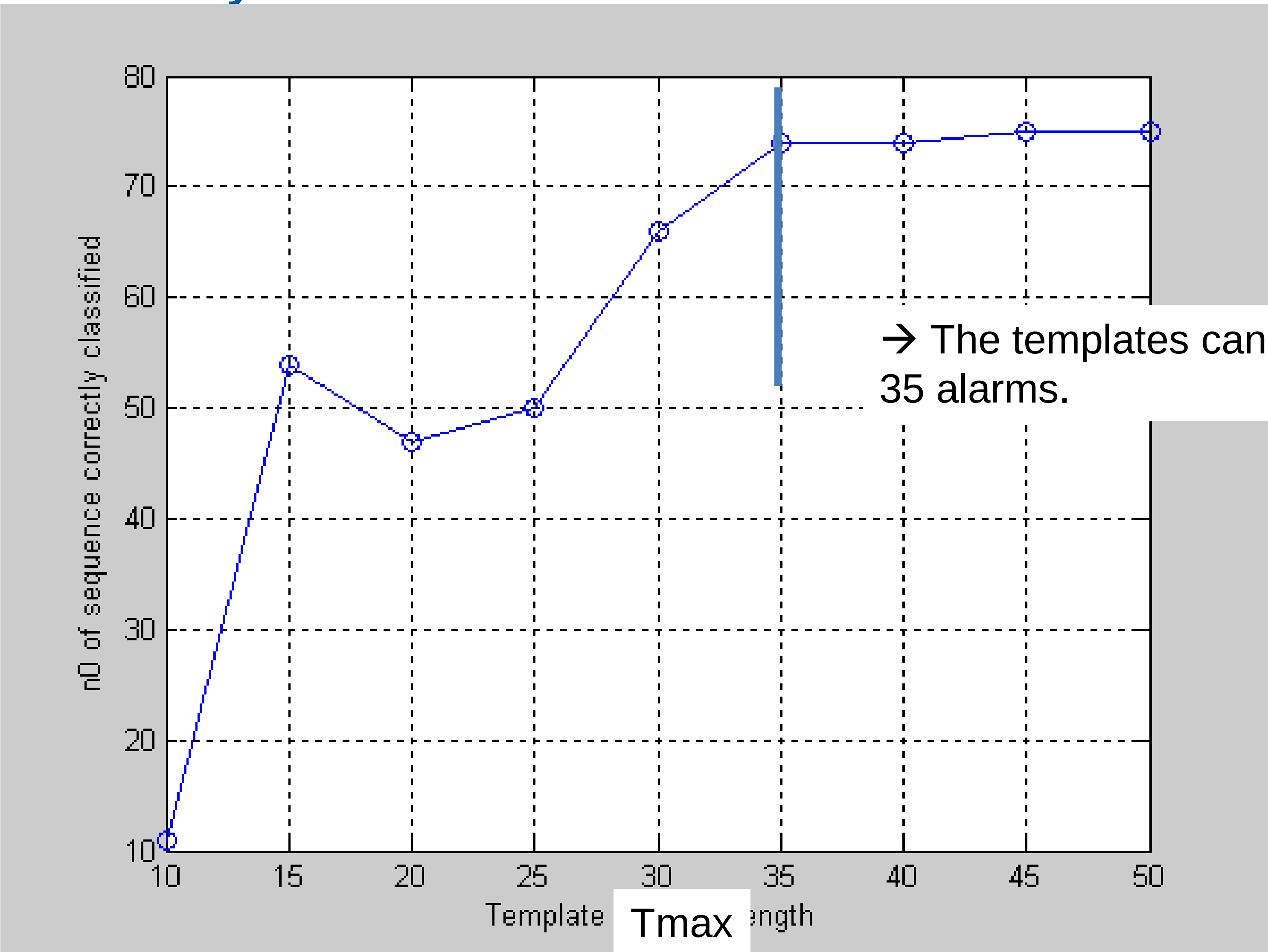
Sequential representation ; weighted similarity :

Full template, full alarm sequence : **75/78**

Impact of the template length



Impact of the template length on the classification accuracy

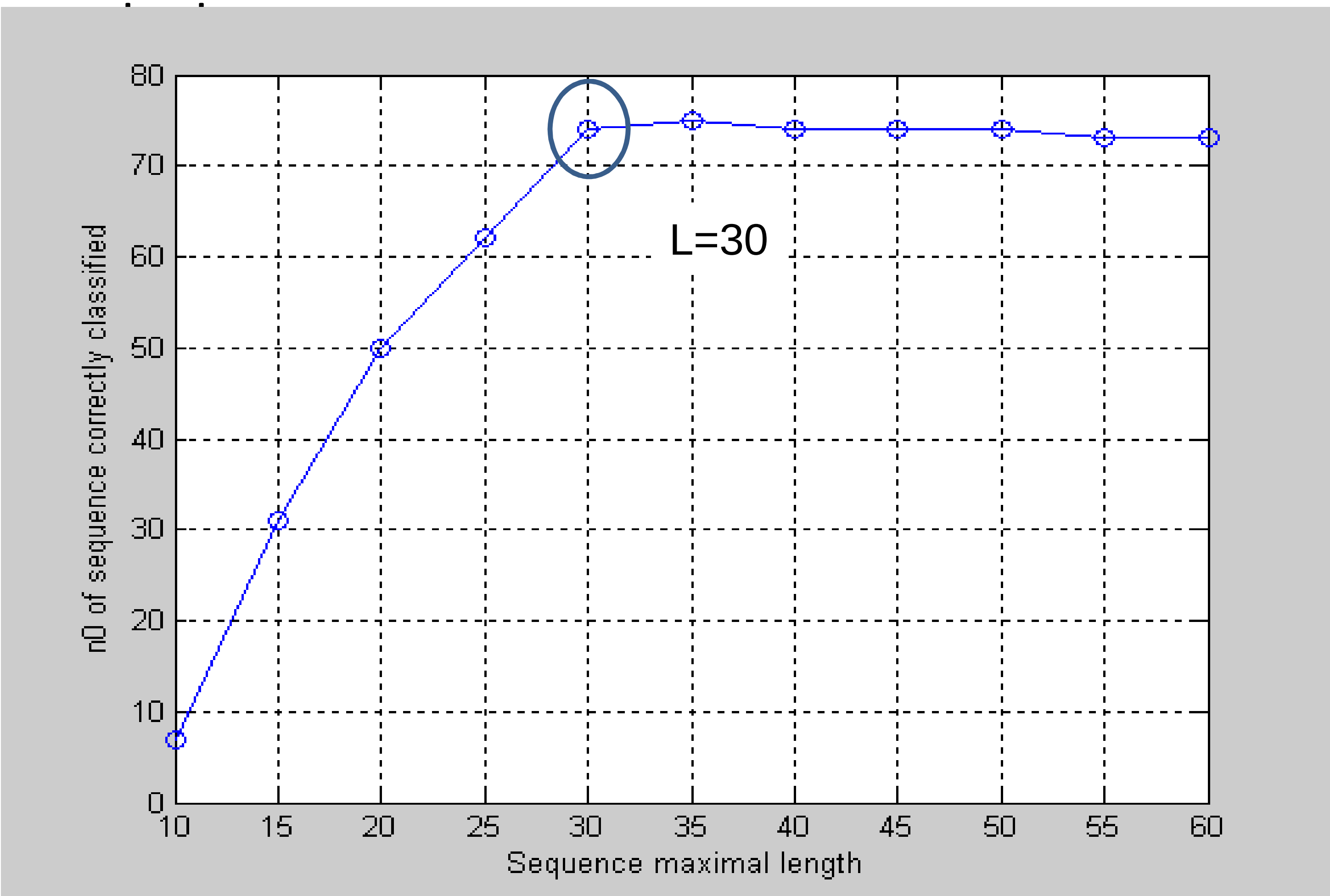


→ The templates can be shorten to 35 alarms.

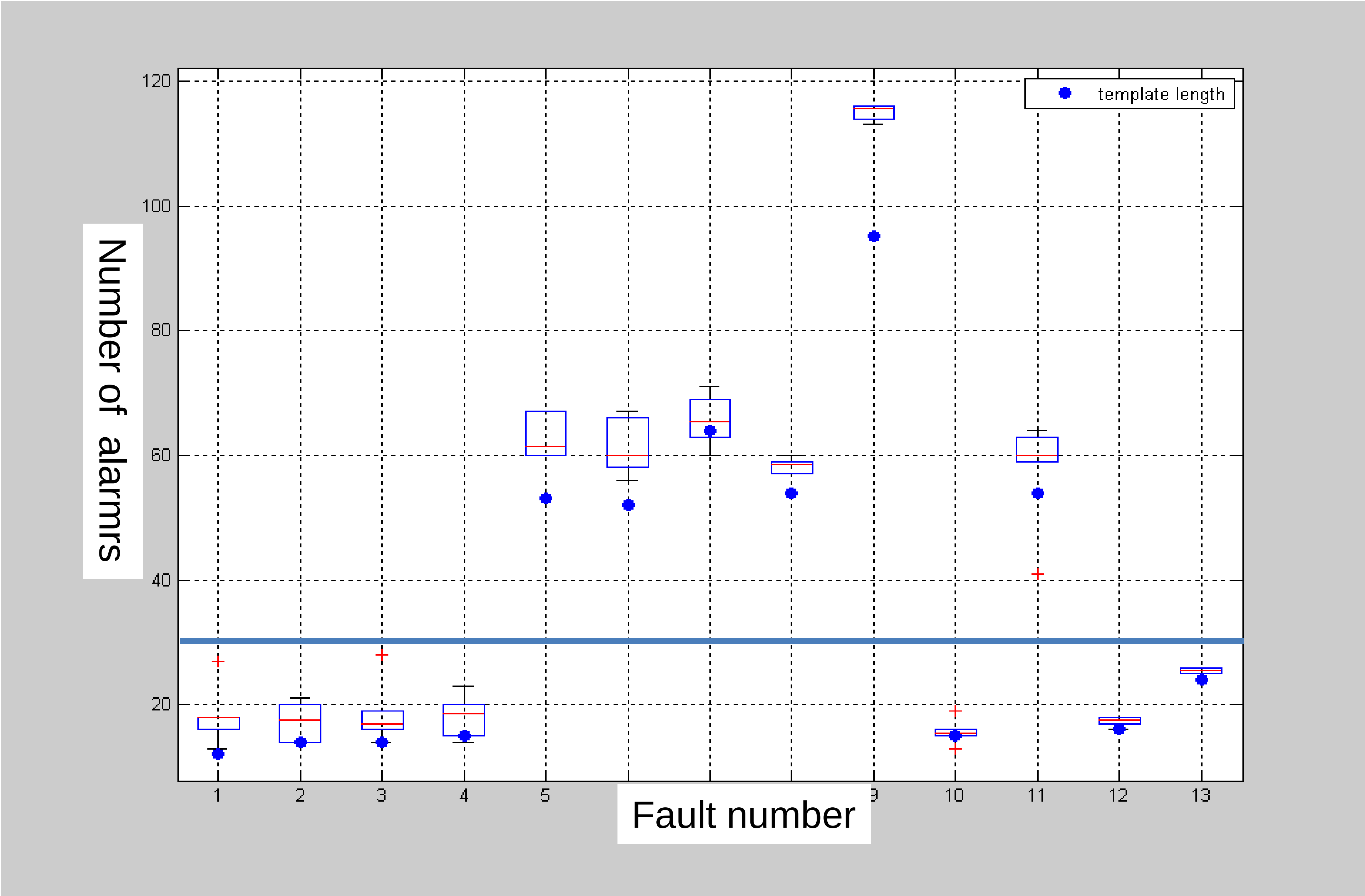
Impact of sequence length

When make the decision? Is it necessary to wait for the end of the alarm flood ?

S is reduced to its L first alarms → the decision is made after L alarms



Impact of sequence length



Conclusion

- **Off line** : A strategy to learn information from a fault data set → valuable feed-back information on the alarm generation process
  - Irrelevant alarms
  - Fault diagnosability from its alarm list
    → Lack of instrumentation
  - Typical fault pattern : guideline for non expert operators
- **On line** : A strategy to diagnose a fault on the occurrence of an alarm flood, based on similarity

➢ No model required but a need for data

➢ Both the presence and the absence of alarms are considered