Investigation of the diagnostic properties of sensors and features in a multiphase flow facility case study

Anna Stief* James R. Ottewill* Jerzy Baranowski**

* ABB Corporate Research Center, ul. Starowiska 13a, 31-038 Krakow, Poland (e-mail: anna.stief@pl.abb.com, james.ottewill@pl.abb.com)

** Department of Automatic Control and Robotics, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, AGH University of Science and Technology, Al. A. Mickiewicza 30, 30-059 Krakow, Poland (e-mail: jb@agh.edu.pl)

Abstract: Feature design and selection is one of the first steps towards successful fault detection and diagnosis. Data from different sources can contain complimentary information about a monitored system. One sensor might be more adept at detecting one fault or operation mode than another sensor. This could be due to differences in the type of the sensor used or simply due to differences in the physical location of two identical sensors relative to the fault source. Hence methods which fuse features from multiple sources can often detect and diagnose a greater number of fault modes with higher confidence. However, solutions that require data from multiple sensors as inputs can be susceptible to failure if one or more of those sensors cease to function. Optimally, a solution will fuse data from a sufficient number of sensors so that the advantages of sensor fusion are realized, while the robustness of the system is retained. In this paper the authors investigate how the best subset of features might differ for fault severity detection and fault diagnosis in a multiphase flow facility case study. ReliefF, which is a K-nearest neighbors-based feature selection filter, is used to rank the features for the different problems. The dataset used for the analysis contains data from various operating conditions and induced faults with various severities. It is shown that the optimal subset of features varied for different monitoring problems. This observation was confirmed by applying a fault tree classifier on the selected subset of features compared to the full set of features.

Keywords: Signal processing, Feature extraction, Feature selection and Process monitoring.

1. INTRODUCTION

Feature design and selection is one of the first steps towards successful fault detection and diagnosis. Data from disparate sources often contains complimentary information about the systems being monitored. One sensor might be more adept at detecting one fault or operation mode than another sensor. This may be due to differences in the type of the sensor used or simply due to differences in the physical location of two identical sensors relative to the fault source. Similarly, one feature derived from a signal recorded from a particular sensor might be capable of detecting one type of fault, while a different feature calculated from the same source might be more successful at detecting a different type of fault. Hence methods which fuse features from multiple sources can often detect and diagnose a greater number of fault modes with higher confidence. Although domain knowledge is a necessity for feature extraction, in complex systems, it might be necessary to rely on additional data-driven methods to reach the best subset of features tailored to a monitoring problem (Onel et al. (2018)). Feature selection aims to reduce the complexity of the system, omit noisy, irrelevant and correlated features, while retaining only the features providing useful information for a monitoring problem (Alkhadafe et al. (2016)). Feature selection can lead to better understanding of the dataset, better learning performance, lower computational cost and more interpretable models (Chandrashekar and Sahin (2014), Miao and Niu (2016)).

Ensuring that a fault detection and diagnosis (FDD) based on sensor fusion is robust against sensor failures is a known challenge (Heng et al. (2009), Jardine et al. (2006)). The system architecture is crucial for successful monitoring (Esteban et al. (2005)). Monitoring systems, which rely on many sensors, are more vulnerable to sensor failures than those, which use fewer sensors. However the observability of fault modes puts a constraint on the minimal number of sensors necessary for the successful diagnosis. A possible solution is to decompose the system into several subsystems for different monitoring problems.
Hence, if a sensor which is not necessary for all of the subsystems fails, only part of the monitoring solution will potentially be compromised. As a result, modular and scalable approaches are preferable, where each monitoring subsystem uses as few sensors as possible, in order to ensure the overall monitoring solution is robust and easy to maintain.

In this paper the authors investigate how the most suitable sensors and subsets of features might differ for different monitoring problems such as fault detection and fault severity diagnosis in a multiphase flow facility case study. High frequency pressure and ultrasonic sensors provide abundant information for flow regime monitoring and fault detection. Power Spectral Density and the Discrete Wavelet Transform are used for feature extraction. ReliefF, which is a K-nearest neighbors-based feature selection filter (Robnik-Šikonja and Kononenko (1997)), is used to rank the features for the different monitoring problems. The analysed dataset contains data from different operating conditions and induced faults with different severities. The optimal subset of features is explored for fault detection and fault severity diagnosis by applying a standard Fault Tree classifier on the selected subset of features compared to the full set of features for the different monitoring problems. An investigation into a multiclass fault tree classifier without feature selection showed how the monitoring accuracy would change for simulated sensor faults.

2. DATASET

The dataset is from a multiphase flow facility case study conducted at the Process System Engineering Laboratory of Cranfield University, (Stief et al. (2018)). The test facility is supplied with water and air, taking the mixed flow through a horizontal section, up to a riser, where the gas and liquid phases are separated and returned in a closed loop. The horizontal and riser sections were instrumented with 9 pressure sensors and an ultrasonic Doppler sensor to monitor the flow regime in the piping, as shown in Figure 1. The pressure sensors were sampled at 5 kHz, while the ultrasonic sensor was sampled at 10 kHz. Three manual valves were used to gradually introduce faults into the process:

- Air leakage: V1 valve was gradually opened to manually induce air leakage on the air input line
- Air blockage: V2 valve was gradually closed to manually induce air blockage on the air input line
- Diverted flow: V3 valve was gradually opened to manually induce diverted flow to the process

The three induced fault scenarios were tested under two operating conditions:

- Operating condition A: 120 Sm$^3$/h air flow rate, 0.1 kg/s water flow rate
- Operating condition B: 150 Sm$^3$/h air flow rate, 0.5 kg/s water flow rate

For each fault and operating condition (OC) the valve positions in degrees are summarised in Table 1. Based on experimental observations, the datasets were classified into three severity categories: Mild, Moderate and Severe.

| Induced fault       | OC   | Mild | Moderate | Severe |
|---------------------|------|------|----------|--------|
| Air leakage         | A    | 5    | 10       | 15     |
|                      | B    | 5    | 10, 15   | 25, 30, 40, 90 |
| Air blockage        | A    | 80, 70, 60 | 50, 40 | 30, 20, 10 |
|                      | B    | 80, 70, 60 | 50, 40 | 30, 20, 10 |
| Diverted flow       | A    | 5, 10, 15 | 20, 30 | 40, 60 |
| flow                | B    | 10, 20, 30 | 30, 40 | 45, 50, 60 |

Table 1. Fault severities by valve openings (°)

Fig. 1. The schematic horizontal and vertical riser with sensors and manual valves

Initially the process was operated at nominally healthy conditions, before the faults were induced by gradually changing the valve positions. The high frequency measurements were recorded when the flow stabilized for a given set point and valve position. The pressure sensors were sampled at 5 kHz, the ultrasonic sensor was sampled at 10 kHz, with both being recorded for a 60-second window for each set point. 49 measurements were stored for operating condition A and B with and without the induced faults.

3. FEATURE EXTRACTION FROM HIGH FREQUENCY MEASUREMENTS FOR TWO-PHASE FLOW MONITORING

As periodic phenomena are common in two-phase flows, frequency domain and time-frequency domain methods provide a good means for analysing oscillation periods in high frequency flow and pressure data (Shang et al. (2004)). The most common feature extraction methods described in the literature for two-phase flow monitoring are based on Power Spectral Density Estimate (PSD) and the Discrete Wavelet Transform (DWT) (Xie et al. (2004)).

3.1 Power Spectral Density

Santoso et al. (2012) used Power Spectral Density (PSD) features, such as average power and variance of power in different frequency bands, from differential pressure data.
where $\overline{s}$ is the spectrum variance can be calculated in the following way:

$$\bar{x} = \frac{\sum f_i P_x(f_i)}{\sum P_x(f_i)} \quad (1)$$

$$\sigma^2_f = \frac{\sum (f_i - \bar{x})^2 P_x(f_i)}{\sum P_x(f_i)} \quad (2)$$

where $\bar{x}$ is the mean spectral power and $P_x$ is the PSD function. The features extracted from the ultrasonic and pressure signals are summarized in Table 2 and Table 3. The Welch method is used to calculate the PSD spectrum (Welch (1967)).

Figure 2 shows the PSD spectrum for the ultrasonic signal for a developing leakage case. It can be observed that for severe leakage ($V1 = 20, 25^\circ$) the pattern is visibly different and for $V1 = 15^\circ$ the spectrum in the lower frequency ranges is distinguishable from the rest, however for milder leakage the signals show similar characteristics as in the case of the normal condition.

### 3.2 Discrete Wavelet Transform

Wavelet analysis is a powerful tool for analysing complex nonlinear signals such as pressure and flow signals present in two-phase flows (Shang et al. (2004)). It has been successfully applied in two-phase flow monitoring to determine the flow regime and bubble sizes (Seleghim Jr and Milloli (2001)). The Discrete Wavelet Transform (DWT) performs convolution of the original signal with a low-pass filter and then subsequently with a high pass filter providing the approximation coefficients and detail coefficients after down-sampling. The transformation can be repeated maximum $\log_2 N$ times if the signal has $N$ data points, producing approximations and detail levels in different frequency bands. The db2 Daubechies wavelet (Malik and Verma (2012)) was chosen to compute the wavelet coefficients. It was successfully applied in the literature for extracting features from ultrasonic signals (Abbagoni and Yeung (2016)) due to its de-noising and energy preservation properties. To extract representative features from wavelet coefficients standard statistical features can be applied, such as mean, variance, minimum value and maximum value. The extracted DWT signals are shown in Table 4 and Table 5. For both signal types the decomposition stopped at Level 7, as the signals contained increased noise in the lower frequencies.

| Feature type    | Frequency range (Hz) | Feature name |
|-----------------|----------------------|-------------|
| Average power   | 0-120                | USB1        |
|                 | 120-240              | USB2        |
|                 | 240-360              | USB3        |
|                 | 360-480              | USB4        |
|                 | 480-600              | USB5        |
| Mean spectral power | 0-2500              | USmsp       |
| Variance of spectral power | 0-2500 | USvsp      |

Table 2. PSD features from the ultrasonic signal

| Feature type    | Frequency range (Hz) | Feature name |
|-----------------|----------------------|-------------|
| Average power   | 0-240                | PXB1        |
|                 | 240-1000             | PXB2        |
|                 | 1000-1500            | PXB3        |
|                 | 1500-2500            | PXB4        |
| Mean spectral power | 0-2500              | PXmsp       |
| Variance of spectral power | 0-2500 | PXvsp      |

Table 3. PSD features from the pressure signals

| Details wavelet coefficients | Frequency range (Hz) | Feature name |
|------------------------------|----------------------|-------------|
| Level 1                      | 2500-5000            | USL1        |
| Level 2                      | 1250-2500            | USL2        |
| Level 3                      | 625-1250             | USL3        |
| Level 4                      | 312-625              | USL4        |
| Level 5                      | 156-312              | USL5        |
| Level 6                      | 78-156               | USL6        |
| Level 7                      | 39-78                | USL7        |

Table 4. DWT features for the Ultrasonic signal

| Details wavelet coefficients | Frequency range (Hz) | Feature name |
|------------------------------|----------------------|-------------|
| Level 1                      | 1250-2500            | PXL1        |
| Level 2                      | 625-1250             | PXL2        |
| Level 3                      | 312-625              | PXL3        |
| Level 4                      | 156-312              | PXL4        |
| Level 5                      | 78-156               | PXL5        |
| Level 6                      | 39-78                | PXL6        |
| Level 7                      | 19-39                | PXL7        |

Table 5. DWT features for the Pressure signal
4. FEATURe SELECTION WITH RELIEFF

Relief is a supervised, multivariate, feature selection filter which calculates relevance indexes for all the features using a nearest neighbor based joint relationship with the classification target. The original Relief ranks the features for two-class classification problems (Kira and Rendell (1992)), while ReliefF extends Relief to multiclass problems (Kononenko (1994)). The Relief family is efficient, reliable and powerful in estimating the quality of attributes (Huang (2015)). It is also relatively simple and computationally cheap, making it attractive as a feature selection choice. The original Relief takes all of the features and class labels for each observation and outputs a relevance index for each of the features. It starts by randomly choosing an observation and for each feature it searches for the nearest neighbor in the same class (nearest hit) and the nearest neighbor in different classes (nearest miss). A relevance index is then calculated based on the Manhattan distance between the chosen and the found observations: greater weights are given to those features which are close to one another in the same class, less weight is given to those features, which are close to one another in the different class. The resulting ranking is based on how well the features differentiate the observations of different classes. The multiclass ReliefF takes the k-nearest neighbors, and calculates the weights for each feature using Equation 3 (Robnik-Šikonja and Kononenko (2003)):

\[
W[f] = W[f] - \sum_{j=1}^{k} \frac{\text{diff}(f, R_i, H_j)}{mk} + \sum_{j=1}^{k} \frac{\text{diff}(f, R_i, M_j)}{mk} \tag{3}
\]

where \( R_i \) is the chosen observation, \( H_j \) are the nearest hits, \( M_j \) are the nearest misses, \( m \) is the number of iterations defined by the user, \( \text{diff} \) is the Manhattan distance and \( f \) denotes a feature. In this paper ReliefF is used as it is more robust against noisy data.

5. IMPLEMENTATION

The 49 measurements were labelled according to health state (Normal, Blockage, Leakage, Diverted flow). Each 60-second measurement window was divided into 1-second windows and features were extracted from each measurement window. Once the 341 features were extracted the data were divided into 70% training set, 30% test set. The observations were relabelled to fit binary-class classification problems: for each condition the data was relabelled in such a way that all the other conditions were labelled as Others. Once the data was prepared, the ReliefF algorithm ranked the training features by relevance for each fault class. A fault tree classifier was chosen for classification, the advantage this methods lies in its simplicity, the default Matlab® implementation of \textit{fitctree} was used with the Standard CART algorithm (Breiman (2017)). The classifier was trained on the training set and tested on the test set. Initially only the most relevant features were included into the model. The training and testing was subsequently repeated multiple times, with each successive model incorporating an additional feature, ordered in terms of ranking, until all of the features were selected. The performance of the classifier was determined by its classification accuracy. The faulty measurements were also labelled according to fault severity (Mild, Moderate, Severe). For each fault class ReliefF ranked the features by relevance to obtain the best class separation for severity. A fault tree classifier was trained and tested in a similar way as the Relieff detection to determine the classification accuracies relative to selected features.

6. RESULTS

6.1 Feature ranking for fault detection

The ReliefF algorithm provided the feature ranking for the three fault classes. The results are presented in Table B. ReliefF was also applied for the multiclass problem for comparison. P5 pressure sensor provides the most relevant features for all three fault classes. Its high ranking, along with the other P6, P7 and P9 sensors on the riser is more robust against noisy data. The ultrasonic sensor is not present in the top 20 features, which indicates the fact that...
the pressure sensors are most sensitive to the induced faults. This might be due to the location of the ultrasonic sensor at the very top of the riser being less sensitive for pressure and flow fluctuations as the other sensors located along the riser. The DWT features are more relevant than the PSD features in a ratio of 4:1, based on the feature ranking results in Table 6. The results of classification for the three fault classes and for the multiclass case are presented in Figure 3, where the accuracy of classification is investigated in comparison to the number of selected features. Accuracy is defined as the ratio of correct predictions from all the observations in the test set. The Fault tree classifier is very robust against irrelevant features, however it can be observed that high accuracy is possible to achieve using only the first 10-20 features, which in the case of all fault cases comes from pressure sensors P5, P6 and P8.

6.2 Sensor failure simulation

In this section the effects of a possible sensor fault on the accuracy of a monitoring result is investigated. Consider the FDD model constructed using data from sensor $S_1$, $S_2$, ..., $S_N$ as shown in Figure 4. The FDD model fuses data in order to provide a monitoring result. In this investigation, the multiclass classifier, results for which were given in the previous section, was considered as the FDD model. Without feature selection the multiclass classifier achieved an accuracy of 90%. To test the robustness of the classifier, a sensor failure was simulated by setting all values in the signal output from a given sensor equal to zero. Such simulation was conducted for all of the sensors. The achieved fault detection accuracies are summarised in Table 7. For those sensor which did not provide relevant features as listed in Table 6, the conclusion as follows: for P1 and P3 there was no drop in the monitoring accuracy, for the ultrasonic sensor and P4 a small performance degradation is visible. However in case of a failure of P2, which is not among sensors providing relevant features for the multiclass case, the accuracy of the classifier dropped to 61.34%. This illustrates the point while incorporating irrelevant features into a FDD model does not necessarily result in performance degradation in the case of a well-parametrized classifier; sensor failures can still influence the performance, even in the case of failure of seemingly irrelevant sensors. For those sensors which provided the most relevant features like P5 and P6, it can be observed that the accuracy dramatically drops, therefore their importance is further confirmed. For P7 and P9 the accuracy slightly drops, while for P8 the accuracy remains 90%. This shows the importance of selecting the optimal number of features from the ranked feature set. In Figure 3 the multiclass classifier has an accuracy of 92.35% by using only the first 8 features, all originating from P5. The overall optimum is at 18 features (92.6%), including P5, P6 and P7. In such situations, a compromise might be considered, retaining fewer sensors in the monitoring system, even if it means a small percentage loss in accuracy.

![Fig. 4. FDD model fusing N sensors](image)

**Fig. 4. FDD model fusing N sensors**

| Sensor | Blockage Severity | Leakage Severity | Diverted Severity |
|--------|------------------|------------------|-------------------|
| P5     | P4               | P3               | P2                |
| P4     | P3               | P2               | P1                |

**Table 7. Fault detection accuracies with simulated sensor faults (in case there was no sensor failure the accuracy was 90%)**

![Fig. 5. Fault severity detection accuracy using the first N features](image)

**Fig. 5. Fault severity detection accuracy using the first N features**

| no. | Blockage | Leakage | Diverted |
|-----|----------|---------|----------|
| 1   | P5       | P5      | P5       |
| 2   | P4       | P4      | P4       |
| 3   | P3       | P3      | P3       |
| 4   | P2       | P2      | P2       |
| 5   | P1       | P1      | P1       |
| 6   | P5       | P5      | P5       |
| 7   | P4       | P4      | P4       |
| 8   | P3       | P3      | P3       |
| 9   | P2       | P2      | P2       |
| 10  | P1       | P1      | P1       |
| 11  | P5       | P5      | P5       |
| 12  | P4       | P4      | P4       |
| 13  | P3       | P3      | P3       |
| 14  | P2       | P2      | P2       |
| 15  | P1       | P1      | P1       |
| 16  | P5       | P5      | P5       |
| 17  | P4       | P4      | P4       |
| 18  | P3       | P3      | P3       |
| 19  | P2       | P2      | P2       |
| 20  | P1       | P1      | P1       |

**Table 8. Feature ranking for fault severity diagnosis**
6.3 Feature ranking for fault severity diagnosis

The ranked features for diagnosing fault severities are presented in Table 8. Those sensors which did not provide relevant features for fault detection, proved to be useful for fault severity detection, as P1, P2, P3, P4 and the ultrasonic sensor appeared among the most important features. The ratio of DWT and PSD features are almost equal in Table 8. For Blockage the sensors on the horizontal pipeline provided the most relevant features, which is due to the fact that those features are closer to the location of V2 valve. The features relevant for diagnosing the severity of leakage are mostly the mean of spectral power and variance of spectral power from the PSD features. All of the top 10 features for leakage severity detection are from sensors located on the riser, which is according to expectations: severe leakage caused slugging in the riser with significant pressure fluctuations. The results of fault severity detection of the three fault classes are presented in Figure 5, confirming that fault severity detection is harder for the fault detection case. Leakage severity seemed to be equal in Table 8. For Blockage the sensors on the horizontal regimes using an ultrasonic doppler sensor and a neural network. Measurement Science and Technology, 27(8), 084002.

Akhradaf, H., Al-Habaibeh, A., and Lotfi, A. (2016). Condition monitoring of helical gears using automated selection of features and sensors. Measurement, 93, 164–177.

Breiman, L. (2017). Classification and regression trees. Routledge.

Chandrashekar, G. and Sahin, F. (2014). A survey on feature selection methods. Computers & Electrical Engineering, 40(1), 16-28.

Esteban, J., Starr, A., Willetts, R., Hannah, P., and Bryanston-Cross, P. (2005). A review of data fusion models and architectures: towards engineering guidelines. Neural Computing & Applications, 14(4), 273–281.

Heng, A., Zhang, S., Tan, A.C., and Mathew, J. (2009). Rotating machinery prognostics: State of the art, challenges and opportunities. Mechanical systems and signal processing, 23(3), 724–739.

Huang, S.H. (2015). Supervised feature selection: A tutorial. Artificial Intelligence Research, 4(2), 22.

Jardine, A.K., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical systems and signal processing, 20(7), 1483–1510.

Kira, K. and Rendell, L.A. (1992). A practical approach to feature selection. In Machine Learning Proceedings 1992, 249–256. Elsevier.

Kononenko, I. (1994). Estimating attributes: analysis and extensions of relief. In European conference on machine learning, 171–182. Springer.

Malik, S. and Verma, V. (2012). Comparative analysis of dct, haar and daubechies wavelet for image compression. International Journal of Applied Engineering Research, 7(11), 2012.

Miao, J. and Niu, L. (2016). A survey on feature selection. Procedia Computer Science, 91, 919–926.

Onel, M., Kieslich, C.A., Guzman, Y.A., Floudas, C.A., and Pistikopoulos, E.N. (2018). Big data approach to batch process monitoring: Simultaneous fault detection and diagnosis using nonlinear support vector machine-based feature selection. Computers & Chemical Engineering, 115, 46–63.

Robnik-Sikonja, M. and Kononenko, I. (1997). An adaptation of relief for attribute estimation in regression. In Machine Learning: Proceedings of the Fourteenth International Conference (ICML’97), volume 5, 296–304.

Robnik-Sikonja, M. and Kononenko, I. (2003). Theoretical and empirical analysis of relief and relieff. Machine learning, 53(1-2), 25–69.

Santoso, B. et al. (2012). The identification of gas-liquid co-current two phase flow pattern in a horizontal pipe using the power spectral density and the artificial neural network (ann). Modern Applied Science, 6(9), 56–67.

Seleghim Jr, P. and Milioli, F.E. (2001). Improving the determination of bubble size histograms by wavelet denoising techniques. Powder Technology, 115(2), 114–123.

Shang, Z., Yang, R., Cao, X., and Yang, Y. (2004). An investigation of two-phase flow instability using wavelet signal extraction technique. Nuclear engineering and design, 232(2), 157–163.

7. CONCLUSION

This paper has investigated the diagnostic properties of sensors and features in a multiphase flow facility case study. ReliefF, which is a supervised feature selection method, was used to evaluate which features were best suited for fault detection and fault severity diagnosis from feature data obtained from several high frequency pressure sensors and an ultrasonic sensor. Feature selection can not only help in identifying relevant features but also in indicating the relevant sensors for monitoring. The best subset of features differed for the different monitoring problems. This observation was confirmed by applying a standard Fault Tree classifier on the selected subset of features compared to the full set of features for different monitoring problems. It was also shown that using a multiclass fault tree classifier without feature selection makes the system less robust against sensor faults. It has been highlighted that while using irrelevant features does not necessarily result in performance degradation in the case of a well-parametrized fault classifier, sensor failures can have a significant influence on monitoring performance, even in case of failure of seemingly irrelevant sensors. Future work will focus on developing an integrated fault detection and fault severity diagnosis framework applying feature selection and developing decision-level fusion methods to exploit the strengths of a modular fault detection approach which utilizes an optimized number of sensors for each monitoring sub-problem.

REFERENCES

Abbagoni, B.M. and Yeung, H. (2016). Non-invasive classification of gas–liquid two-phase horizontal flow

conditions persisted.
Stief, A., Tan, R., Cao, Y., and Ottewill, J.R. (2018). Analytics of heterogeneous process data: Multiphase flow facility case study. In 10th IFAC Symposium on Advanced Control of Chemical Processes.

Sun, Z. and Zhang, H. (2007). Neural networks approach for prediction of gas-liquid two-phase flow pattern based on frequency domain analysis of vortex flowmeter signals. Measurement Science and Technology, 19(1), 015401.

Welch, P. (1967). The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Transactions on audio and electroacoustics, 15(2), 70–73.

Xie, T., Ghiaasiaan, S., and Karrila, S. (2004). Artificial neural network approach for flow regime classification in gas–liquid–fiber flows based on frequency domain analysis of pressure signals. Chemical Engineering Science, 59(11), 2241–2251.