

# Fault Diagnosis of Heavy Duty Machines: Automatic Transmission Clutches

Tomas Olsson<sup>1,3</sup>, Elisabeth Källström<sup>2</sup>, Daniel Gillblad<sup>3</sup>, Peter Funk<sup>1</sup>, John Lindström<sup>4</sup>, Lars Håkansson<sup>5</sup>, Joakim Lundin<sup>2</sup>, Magnus Svensson<sup>6</sup>,  
Jonas Larsson<sup>2</sup>

<sup>1</sup> School of Innovation, Design, and Engineering, Mälardalen University, Västerås, Sweden

`{tomas.olsson,peter.funk}@mdh.se`

<sup>2</sup> Volvo Construction Equipment, SE-63185 Eskilstuna, Sweden

`{elisabeth.kallstrom, joakim.jl.lundin, jonas.jl.larsson}@volvo.com`

<sup>3</sup> SICS Swedish ICT, Isafjordsgatan 22, Box 1263, SE-164 29 Kista, Sweden

`{tomas.olsson,daniel.gillblad}@sics.se`

<sup>4</sup> ProcessIT Innovations, Luleå University of Technology, Luleå, Sweden

`john.lindstrom@ltu.se`

<sup>5</sup> Department of Applied Signal Processing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden

`lars.hakansson@bth.se`

<sup>6</sup> Volvo Group Trucks Technology, Götaverksgatan 10, 405 08 Göteborg, Sweden

`magnus.svensson@volvo.com`

**Abstract.** This paper presents a generic approach to fault diagnosis of heavy duty machines that combines signal processing, statistics, machine learning, and case-based reasoning for on-board and off-board analysis. The used methods complement each other in that the on-board methods are fast and light-weight, while case-based reasoning is used off-board for fault diagnosis and for retrieving cases as support in manual decision making. Three major contributions are novel approaches to detecting clutch slippage, anomaly detection, and case-based diagnosis that is closely integrated with the anomaly detection model. As example application, the proposed approach has been applied to diagnosing the root cause of clutch slippage in automatic transmissions.

**Keywords:** Case-based Reasoning, Machine Learning, Signal Processing, Fault Diagnosis

## 1 Introduction

Many heavy duty machines, like construction machines, are complex vehicles with many interdependent parts. As time goes by, new features are added and thereby, the complexity is further increased. As a consequence, this leads to longer time in diagnosing faults and often, unnecessary replacement of parts to address the source of problem.

In order to reduce above problems, we are developing an automated, data driven approach to fault diagnosis that in a novel way combines methods from signal processing, statistics, machine learning and case-based reasoning (CBR). The system will continuously monitor a vehicle on-board to detect potential problems (anomalies) and then classify the problem off-board, either using CBR or manually with the help of experts.

CBR has been used for fault diagnosis since the beginning of the CBR field [1-3], and often it has been combined with methods from signal processing and machine learning [4-6]. In the proposed approach, the different methods complement each other in that signal processing and machine learning provide fast and light-weight on-board feature extraction and fault detection, while CBR, in addition to fault diagnosis, can retrieve relevant cases as decision support [7]. In addition, a major contribution, apart from a novel solution to fault diagnosis, is to integrate CBR and the statistical anomaly detection model using a theoretically sound approach to defining similarity grounded in statistics and information theory [8,9]. Previously, it was used as a complement to probabilistic predictions, while it is central for the current work. Furthermore, as an example application, the approach was applied to identifying the root cause of clutch slippage, and therefore, a method for detecting slipping clutches is proposed.

The rest of the paper is organised as follows. Section 2 presents the example problem domain and the used statistical methods. Section 3 gives an overview of the system. Section 4 describes how data was collected and how signal processing is used for extracting features. Section 5 presents an approach for detecting faults (anomalies) on-board a machine. Section 6 describes the off-board diagnosis using CBR. Section 7 relates the proposed approach to previous work. Finally, Sect. 8 ends with some conclusions and future work.

## 2 Preliminaries

In this section, we firstly present the problem domain that we are addressing, that is, diagnosis of automatic transmission clutches, considering especially the problem of clutch slippage. Last, we present two statistical machine learning algorithms that we use for anomaly detection.

### 2.1 Automatic Transmission Clutch Slippage

The automatic transmission clutches in wheel loaders are a crucial component of the driveline. A clutch enables connection and transfer of torque between two rotating shafts when engaged [10]. Multiple disc wet clutches are generally used in automatic transmissions. A multiple disc wet clutch pack consists of separator discs, friction discs, lubricant, piston and two shafts. To engage the clutch, a hydraulic induced normal force is applied to the clutch piston thereby clamping together the friction disc and the separator disc, which allows torque transfer between the two shafts [11]. During gear change both the disengaging clutch for previous gear (off-going clutch) and the engaging clutch for next gear

(on-going clutch) have different angular speeds, the engaging clutch’s angular speed drops to zero at the end of engagement due to generated friction [12].

In this paper, we use the problem of clutch slippage as an example application. Clutch slippage occurs when a gear change takes longer time than expected. This can happen due to several reasons, such as, special working conditions and different types of faults. One such root cause is pressure drop in the transmission due to oil leakage. Thus, when the pressure in the transmission is too low, the time to go from one gear to the next will increase. Occurrence of clutch slippage can be detected but the root cause is not that easily identified. Thus, in this work we have started to automate the root cause analysis of clutch slippage as an example use of the proposed approach.

## 2.2 The Gaussian Mixture Model and Logistic Regression

The Gaussian mixture model assumes that cases are generated from a set of clusters modelled as normal (or Gaussian) probability distributions [13]. So, for cases that are numerical vectors of length  $K$ , assuming a mixture with  $Z$  clusters:  $p(\mathbf{x}) = \sum_{z=1}^Z p(\mathbf{x}|z)p_z$  where  $\mathbf{x}$  is a case,  $z$  is a cluster,  $p_z$  is the probability of a cluster and  $p(\mathbf{x}|z)$  is the normally distributed likelihood of  $\mathbf{x}$  conditioned on  $z$ .

Logistic regression is a classifier that can be trained to distinguish between two classes [13]. For the two classes  $c \in \{0, 1\}$  we have the following probability distributions given a feature vector (a case)  $\mathbf{x}$ :  $p(c = 1|\mathbf{x}) = \frac{1}{1+\exp(-\boldsymbol{\omega}^T \mathbf{x})}$  and  $p(c = 0|\mathbf{x}) = \frac{\exp(-\boldsymbol{\omega}^T \mathbf{x})}{1+\exp(-\boldsymbol{\omega}^T \mathbf{x})}$  where  $\boldsymbol{\omega}$  is a weight vector with  $K + 1$  weights assuming that  $x$  has  $K + 1$  features including an extra feature that is 1 for all cases. Then, the case is classified as  $c = 1$  if  $\boldsymbol{\omega}^T \mathbf{x} \geq 0$  and  $c = 0$  otherwise.

## 3 System Overview

In this section, we give an overview of the system for diagnosing wheel loaders where computation is being performed both on-board and off-board the machine. On-board the system filters signals for interesting events that then are analysed off-board. In the future, more automatic diagnosis might be performed on-board, but in the current approach, an important function is to enable decision support to an expert panel for querying the off-board system for relevant cases.

Figure 1 presents an overview of the proposed system, showing the flow of computation from on-board a wheel loader to a central off-board server. The system works as follows (numbers in Fig. 1): (1) The on-board feature extraction component continuously monitors the machine using on-board sensors for interesting events. (2) Extracted features are assessed for anomalies by a combination of a statistical model of the normality of each feature and logistic regression trained on both normal and anomalous data. (3) If classified as anomalous, the case-based diagnosis will classify the fault or indicate whether it is a new type of fault, and assess its severity. Feature extraction will be presented in Sect. 4, anomaly detection in Sect. 5, and case-based diagnosis in Sect. 6. A prototype was implemented using the scikit-learn library [14] and the minepy library [15].

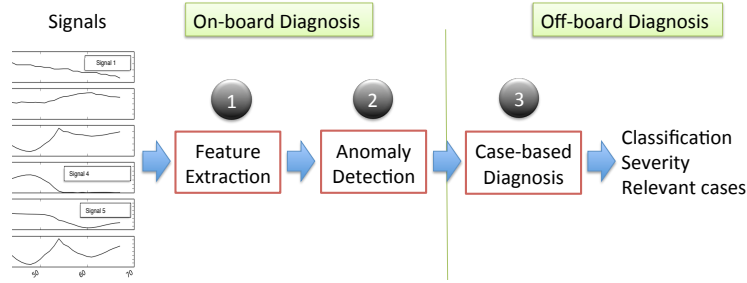


Fig. 1. On-board and off-board fault diagnosis system.

## 4 Data and Extracted Features

This section describes the data collected from the wheel loader and how features are extracted. In this paper, we only consider the gear change between two gears, from clutch 1 to clutch 2. For the experiments that follows, we have used six signals logged from the on-board electronic system (the CAN-bus). The logged signals are the turbine torque, clutch 1 differential speed, clutch 2 differential speed, out-going speed, input speed, and turbine speed. The data was read at a sampling frequency of 500Hz. Collecting non-anomalous signals was easy but in order to get faulty data we had to simulate a fault. Thus, we simulated oil leakage in a clutch by installing two manual needle valves on the pressure out-takes on clutch 1 and clutch 2. Thereby, we could adjust the oil pressure going to the piston in the clutch. The valve can then be opened in seven steps from fully closed to fully opened (0-7 turns) that also simulate the severity of the fault. From the six CAN-bus signals, we have extracted five feature for each gear change: the length in time, mean value, standard deviation, kurtosis and maximum of the sliding mean square value filtering (SMSVF) of the clutch 1 differential speed signal. A clutch slippage is an increase in time of a gear change, so the length of a gear change is clearly relevant. Also the shape of the signal is important, which is captured by the next three features. In experiments, we also validated that computing maximum of the SMSVF gives a good indication when a clutch slippage occurs. However, it is not in itself enough for identifying the root cause. Below we describe the SMSVF and kurtosis in more details.

### 4.1 Mean Square Value and Sliding Mean Square Value filtering

The mean square value of a signal is related to the power of the signal [16]. If a sampled signal is weakly ergodic, an estimate of its mean square value may be calculated by squaring each signal sample and sum them, and finally dividing this sum with the number of samples in the sum [16]. The mean square value of a weakly ergodic signal  $x(n)$ ,  $n = 1, 2, \dots, N$  is given as

$$\bar{x}^2 = \frac{1}{N} \sum_{n=1}^N x(n)^2 \quad (1)$$

Where  $N$  is equal to the number of samples in the average. The Sliding mean Square Value filtering is realised by filtering the square of a sample of a signal with an adequate filter [17]. The moving Average filter is frequently used as a de-noising technique because of its simplicity in implementation and low computational load [18]. The mean square value estimates may also provide information about the stationarity of a signal [19].

## 4.2 Higher-order cumulant: Kurtosis

Higher order cumulants has been trending in diverse applications for many years for their ability to handle non-Gaussian processes [20]. Cumulants above the third-order are regarded as higher order cumulants while lower order cumulants are from the third-order and below [21]. Higher-order cumulants are preferred instead of second-order for signals corrupted with Gaussian measurement noise since they are blind to Gaussian processes [20]. The first order cumulant is the mean value, while the second order cumulant is the variance and the third order cumulant is the third central moment or skewness [20, 21]. Kurtosis is based on the fourth order cumulant and thus it is a higher order cumulant [22]. The kurtosis gives an indication of the peakedness and tailedness of a distribution [22]. The Kurtosis is the normalized fourth order cumulant about the mean and it is expressed as  $Kurtosis = E[(x(t) - E[x(t)])^4] / ((E[(x(t) - E[x(t)])^2])^2) = \mu_4 / \sigma^4$  Where  $\mu_4$  is the fourth order cumulant and  $\sigma$  is the standard deviation [22].

## 5 Anomaly Detection

This section presents our work in developing on-board anomaly detection that uses the output of the feature extraction component. We assume that there is a very large set of cases known to be normal and a relatively small set of cases known to be anomalous. In addition, not all fault classes are known beforehand so new faults should also be detected. Thus, given these assumptions, an ordinary classifier is not sufficient. So, we use an anomaly detection approach instead.

A common way of doing anomaly detection is to fit a statistical model to the non-anomalous cases and then, by choosing a suitable threshold, classify cases above the threshold as normal and below the threshold as anomalous since they are unlikely [23]. In the following, we instead consider this as a binary classification problem, so that a soft threshold can be introduced by training a probabilistic classifier using as case features the output of the statistical model.

The anomaly detection component consists of two parts: the first part is the statistical modelling part where we fit GMMs to the non-anomalous data and the second is the classification part where we fit logistic regression to distinguish between anomalous and non-anomalous data. In most cases when doing anomaly detection, a GMM is fitted directly to the whole feature vectors, while we instead fit a GMM to each feature in the feature vectors, independently of the other features. The output from the GMMs is then a new feature vector with the log-likelihood of each pair of feature value and cluster. For instance, if we have 5

signals for which we extract 5 features each that result in 25 features in total and then run a GMM using 5 clusters each, then we end up with feature vector of length 125. The GMM step can be considered as making an anomaly score for each pair of feature and cluster that are then fused together in the next step using logistic regression. In other words, the linear regression will learn the weights and thresholds for distinguishing between normal and anomalous signals from the log-likelihood features.

For the experiment, we trained the GMM and logistic regression on 80% normal data and 50% fault data and the rest of the data was used for testing and validation. This was repeated 10 times and the resulting AUC measures were then averaged. The total size of the data set is 337 of which 63 cases are faults. Table 1 shows the results measured in the Area Under the Curve (AUC) that is a common performance measure for anomaly detection based on the receiver operating characteristic curves (ROC curves) [15]. The AUC is a value between 0.5 and 1, where 1 means perfect detection while 0.5 means completely random detection. The two AUC columns compare the use of only the mean,

**Table 1.** The Area Under the Curve (AUC) for the anomaly detection.

Feature Type	AUC (3 features)	AUC (5 features)
Original features	0.948	0.965
Auto cluster	0.959	0.949
Clusters (12 and 14))	0.968	0.955

standard deviation and the length of the signal as original features (3 features) and, in addition to that, the kurtosis and the squared mean value (5 features). In the table, *Original features* means that we do not use the GMM but the original features, while *Auto cluster* means that we use an automatic means of selecting the number of clusters for each feature that evaluate the goodness of fit of the statistical model and *Clusters* means that we used the same number of clusters for all features but selected the number of clusters that had the best validation AUC. As can be seen in the table, the GMM (the last two rows) does not improve the AUC, which is already quite high. Thus, we are able to learn to detect faults not far from perfect (1.0), but the use of GMM does not improve the performance.

## 6 Case-Based Diagnosis

In this section, we describe our approach to diagnosis using CBR that integrate the statistical model from previous section. An essential component of a CBR approach is a similarity metric that measures the usefulness of a case compared to a new case. In our approach, we measure similarity between cases as how similarly they deviate from the normal cases with respect to the statistical anomaly detection model from Sect. 5, and thereby, integrates CBR with the GMMs.

In [9], we defined the similarity between two cases using an information theoretical metric of similarity – the symmetric Kullback-Leibler divergence (J-divergence) – that compared the predicted class distributions of the two cases [24]. Likewise, in this paper, we derive a similarity metric for comparing cases with respect to the cluster distributions of the GMMs, which however is a multi-label problem in that each case can belong to many clusters. Thus, we let  $\mathbf{z}$  be a vector with one cluster  $z^k \in \{1, \dots, Z\}$  for each feature  $k = 1, \dots, K$ , then the J-divergence between two cases  $\mathbf{x}_i, \mathbf{x}_j$  relative to the distribution of  $\mathbf{z}$  is

$$\begin{aligned} J(\mathbf{x}_i, \mathbf{x}_j) &= \sum_{\mathbf{z}} (\log(p(\mathbf{z}|\mathbf{x}_i)) - \log(p(\mathbf{z}|\mathbf{x}_j))) (p(\mathbf{z}|\mathbf{x}_i) - p(\mathbf{z}|\mathbf{x}_j)) \\ &= \sum_{k=1}^K \sum_{z^k=1}^Z (\log(p(z^k|x_i^k)) - \log(p(z^k|x_j^k))) (p(z^k|x_i^k) - p(z^k|x_j^k)) \\ &\leq \sum_{k=1}^K \sum_{z^k=1}^Z |\log(p(x_i^k|z^k)) - \log(p(x_j^k|z^k))| \end{aligned}$$

where  $p(\mathbf{z}|\mathbf{x}_i) = p(z^1|x_i^1) \cdot p(z^2|x_i^2) \cdot \dots \cdot p(z^K|x_i^K)$  (assuming independence of cluster  $z^k$  given feature  $k$  and value  $x_i^k$ ) and  $p(z^k|x^k) = \frac{p(x^k|z^k)p_{z^k}}{p(x^k)}$ . The less-than-equal row is valid since  $\log(p(z^k|x_i^k)) - \log(p(z^k|x_j^k)) = \log(p(x_i^k|z^k)) - \log(p(x_j^k|z^k)) + \log(p(x_j^k)) - \log(p(x_i^k))$  (last terms are independent of  $z^k$  and thereby canceled out in the sum) and  $\max(|p(z^k|x_i^k) - p(z^k|x_j^k)|) = 1$ . Thus, starting from the J-divergence, we can derive the Manhattan distance for the log-likelihood feature vectors that were defined in the previous section.

For evaluation, we use the same data as for anomaly detection. However, since we only have data from one type of fault, but with varying severity, we only evaluate the predicted severity. Thus, we train the k-nearest neighbour algorithm to predict the valve opening using the average of the most similar cases. In the experiments, we performed 5 times leave-two-out cross-validation using one case for testing and one case for validation, and the remaining for training. The features were normalised to have a mean of 0 and a standard deviation of 1. Then, each feature was weighted using the maximum information coefficient (MIC) between the predicted attribute and the features [13]. The number of neighbours was selected using 5-fold cross validation. The result is shown in Table 2. For both the original features and log-likelihood features, we

**Table 2.** The mean squared error (MSE) for clutch slip severity diagnosis.

Feature Type	MSE (3 features)	MSE (5 features)
Original features	4.32	3.91
Auto cluster	4.59	5.24
Clusters (14 and 5))	4.36	5.18

used the Manhattan distance. As can be seen, the combination of 5 features and Original features has the lowest mean squared error, although the performance is not very good. The clustering seems not to have any positive effect in this case either.

## 7 Related Work

There are several fault diagnosis approaches that combine CBR with other methods. One application in the INRECA project is fault diagnosis of robots that integrates causal trees, decision trees and CBR [4]. A hybrid CBR system with an ART-Kohonen neural network (ART-KNN) for diagnosing an electric engine is described in [5]. A CBR approach for diagnosing faulty robot gearboxes was presented in [25, 26] that uses methods from signal processing, the Discrete Wavelet Transform as well as the Discrete and Fast Fourier Transform. In [6], an advanced CBR system for automobile service troubleshooting is described that integrates the use of associate-rule mining, CBR and text mining. In [27], the authors describe an approach resembling ours for fault detection in locomotives that is an add on to the CBR diagnosis system ICAROS presented in [28]. Like in our system, the signals are processed individually to detect an anomaly and then fused together using another machine learning algorithm. However, the integration of the CBR and anomaly detection systems is not described.

## 8 Conclusions and Future Work

This paper presents a novel approach to fault diagnosis in heavy duty vehicles that was applied to diagnosing the root cause of clutch slippage of automatic transmissions. The approach integrates methods from signal processing, statistics, machine learning and CBR. We have investigated five different types of extracted features, among which one feature is a novel approach to detecting clutch slippage. We have also presented an approach to anomaly detection that combines GMMs with logistic regression. In addition, we have defined similarity metrics that integrates the case-based diagnosis with the statistical anomaly detection model in a novel way.

In addition, we have reported preliminary results showing that the approach works in that it is able to learn from the original features, but more data is needed, including additional fault types, in order to draw more precise conclusions. Another issue is that the anomaly detection and CBR were evaluated in isolation, while they would be connected in a real system. Thus, a normal case would have a valve opening of 0 if also evaluated by for its severity. Also, any false positives generated by the anomaly detector should be considered, so that the CBR would be able to classify them as normal.

One possible use of the proposed approach is to support emerging business models which use monitoring for predicting problems to prevent risk due to failure, such emerging business models includes, Product-Service Systems, Industrial Product-Service Systems and Functional Products [29]. Yet, there is a



lot of work left before this becomes a reality. We could further investigate other types of features for improved performance. The anomaly detection and logistic regression could also be compared with other related approaches. The case-based diagnosis could use other metrics or be compared to other learning approaches.

## Acknowledgements

This work has been partially supported by the FP7 EU Largescale Integrating Project SMART VORTEX co-financed by the European Union [30], the Swedish Knowledge Foundation (KK-stiftelsen) [31] through ITS-EASY Research School and Swedish Governmental Agency for Innovation Systems (VINNOVA) grant no 10020 and JU grant no 100266.

## References

1. Thomas, R., Thomas, G., Littler, J.: The cognitive role of an engineer in a diagnostic task. In: UK IT 1990 Conference, IET (1990) 259–263
2. Kockskamper, K., Traphoner, R., Wernicke, W., Faupel, B.: Knowledge acquisition in the domain of cnc-machining centers: the moltke approach2. In: EKAW 89: Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems, Paris, July 89, AFCET (1989) 180
3. Althoff, K., Maurer, F., Wess, S., Traphöner, R.: Moltke: An integrated workbench for fault diagnosis in engineering systems. In: Proceeding EXPERSYS-92, Paris (1992)
4. Auriol, E., Crowder, R., McKendrick, R., Rowe, R., Knudsen, T.: Integrating case-based reasoning and hypermedia documentation: an application for the diagnosis of a welding robot at odense steel shipyard. *Engineering Applications of Artificial Intelligence* **12**(6) (1999) 691–703
5. Yang, B., Han, T., Kim, Y.: Integration of art-kohonen neural network and case-based reasoning for intelligent fault diagnosis. *Expert Systems with Applications* **26**(3) (2004) 387–395
6. Chougule, R., Rajpathak, D., Bandyopadhyay, P.: An integrated framework for effective service and repair in the automotive domain: An application of association mining and case-based-reasoning. *Computers in Industry* **62**(7) (2011) 742–754
7. Leake, D., McSherry, D.: Introduction to the special issue on explanation in case-based reasoning. *Artificial Intelligence Review* **24**(2) (2005) 103–108
8. Olsson, T., Gillblad, D., Funk, P., Xiong, N.: Case-based reasoning for explaining probabilistic machine learning. *International journal of computer science & information technology (IJCSIT)* **6**(2) (April 2014)
9. Olsson, T., Gillblad, D., Funk, P., Xiong, N.: Explaining probabilistic fault diagnosis and classification using case-based reasoning. In: *Case-Based Reasoning Research and Development. International Conference on Case-Based Reasoning (ICCBR-2014)*. (2014)
10. Lingesten, N.: Wear behavior of wet clutches. Licentiate thesis / Luleå University of Technology. Luleå tekniska universitet (2012)
11. Mäki, R.: Wet clutch tribology: friction characteristics in limited slip differentials. (2005)

12. Fatima, N., Marklund, P., Larsson, R.: Water contamination effect in wet clutch system. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* **227**(3) (2013) 376–389
13. Murphy, K.P.: *Machine learning: a probabilistic perspective*. MIT Press (2012)
14. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: *Scikit-learn: Machine learning in Python*. *Journal of Machine Learning Research* **12** (2011) 2825–2830
15. Albanese, D., Filosi, M., Visintainer, R., Riccadonna, S., Jurman, G., Furlanello, C.: *minerva and minepy: a c engine for the mine suite and its r, python and matlab wrappers*. *Bioinformatics* **29**(3) (2013) 407–408
16. Campbell, D.: *Statistical Measures of Signals*. School of Computing, University of Paisle
17. Andrén, L., Håkansson, L., Brandt, A., Claesson, I.: Identification of dynamic properties of boring bar vibrations in a continuous boring operation. *Mechanical systems and signal processing* **18**(4) (2004) 869–901
18. Koswatta, R., Karmakar, N.C.: Moving average filtering technique for signal processing in digital section of uwb chipless rfid reader. In: *Asia-Pacific Microwave Conference, IEEE* (2010) 1304–1307
19. Bendat, J.S., Piersol, A.G.: *Random data: analysis and measurement procedures*. Volume 729. John Wiley & Sons (2010)
20. Mendel, J.M.: Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE* **79**(3) (1991) 278–305
21. Cumulant. <http://en.wikipedia.org/wiki/Cumulant> (Last Accessed: July 2014)
22. DeCarlo, L.T.: On the meaning and use of kurtosis. *Psychological methods* **2**(3) (1997) 292
23. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. *ACM Computing Surveys (CSUR)* **41**(3) (2009) 15
24. Kullback, S., Leibler, R.A.: On information and sufficiency. *The Annals of Mathematical Statistics* **22**(1) (1951) 79–86
25. Olsson, E., Funk, P., Xiong, N.: Fault diagnosis in industry using sensor readings and case-based reasoning. *Journal of Intelligent and Fuzzy systems* **15**(1) (2004) 41–46
26. Olsson, E., Funk, P., Bengtsson, M.: Fault diagnosis of industrial robots using acoustic signals and case-based reasoning. In: *Advances in Case-Based Reasoning*. Springer (2004) 686–701
27. Xue, F., Yan, W., Roddy, N., Varma, A.: Operational data based anomaly detection for locomotive diagnostics. In: *MLMTA*. (2006) 236–241
28. Varma, A., Roddy, N.: Icarus: design and deployment of a case-based reasoning system for locomotive diagnostics. *Engineering Applications of Artificial Intelligence* **12**(6) (1999) 681–690
29. Lindström, J., Plankina, D., Nilsson, K., Parida, V., Ylinenpää, H., Karlsson, L.: Functional products: Business model elements. In: *Proceedings of the 5th CIRP International Conference on Industrial Product-Service Systems*. (March 2013)
30. SMART VORTEX: Scalable semantic product data stream management for collaboration and decision making in engineering. <http://www.smartvortex.eu/> (Last Accessed: July 2014)
31. KK-Stiftelse: Swedish Knowledge Foundation. <http://www.kks.se> (Last Accessed: September 2013)