Monitoring of automotive multistage mechanical transmissions using multi-class support vector machine

Mahmoud M.A. Sayed1 and Wessam. M. Hussein2

1Mechanical Department, Canadian International College (CIC), 2Department of Mechanical Engineering, The Military Technical College, Cairo, Egypt

Key words: Condition based maintenance, health monitoring system, Mechanical transmission, support Vector Machine, Vibration analysis

Abstract
The need for semi-autonomous or autonomous operations, communication delay, short contact periods as well as the need for survival in harsh environments poses unique challenges to Automotive Mechanical Transmission Systems (AMTS). Predictive health monitoring (PHM) systems are currently gaining in popularity due to their effectiveness in providing robust information about the system condition and reducing maintenance costs. This paper presents a PHM system for monitoring different gear faults using vibration analysis and Support Vector Machine (SVM) algorithms. Experiments were conducted on a multi-stage gearbox (Automotive Mechanical Transmission Systems) under three conditions, normal, external vibrational excitation and oiling system high temperature. Multi-class SVM based on developing a model for normal and faulty states; the model used for monitoring the upcoming sensory data, and classifies them as normal or faulty ones. The model is verified through additional experimental observations. The classifier algorithm was coded in Matlab and showed a good potential in classifying different failure mechanisms.

I. INTRODUCTION
Monitoring the condition of the in-service mechanical transmission system is an important issue for reliability, where their components deteriorate over the time and affected much when subjected to varying loads. This led in continuous improvement of maintenance strategies from breakdown and periodic maintenance to Condition Based Maintenance (CBM) and predictive maintenance in order to sustain reliability and reducing the periodic maintenance costs. Also, in some applications there is more demanding aspect such as saving man's life other than reliability[1]. Smith[2], has defined the causes of transmission vibration and its transmission path, including factors such as manufacturing error, design error and gear tooth deflection, which combine to introduce a Transmission Error (TE), which is the primary source of the vibration.

Over the past decade, vibration analysis proved to be a trustworthy diagnostic technique that can provide reliable information. However, in the last 10 years researchers devoted a much effort to support CBM actions using vibration information[9-13].

Many Researchers focused on developing multi sensors fusion algorithms to fuse vibration analysis information with other sensory data, such as Acoustic Emission (AE) and oil debris analysis to minimize false alarms that may occur in failure prediction.

Parametric methods based on mathematical modeling is used to fit measured time series waveform data to a parametric time series model, and then extract features based on this model. Two models are currently in use: the auto regressive (AR) and auto-regressive moving average (ARMA) models. The advantage of mathematical modeling based on parametric methods over the neural networks model-based method is its ability to deal with time series data directly without the need for a signal pre-processing step to extract useful features that can be modeled to represent the system. However, they can only be used to model a time series signal such as a vibration signal, and cannot be applied to combined information from several techniques (vibration and AE) such as in the case of fuzzy logic [14-17]. Also, other researchers devoted efforts to build intelligent algorithms based on vibration features including Expert systems, Artificial Neural Network (ANN's), Genetic algorithm, and fuzzy logic [18-29].

Onsy et al.[30-32] is devoting their efforts in developing smart CBM systems that can use one analysis technique only such as vibration or acoustic emission analysis along
with intelligent algorithms to predict the onset of failures; this is to reduce costs of different sensory requirements.

The simplicity of data driven modeling approach is that there is no need for a fundamental model of the system and only data from normal operation needs to be used, which is generally available in some form for most machines. Among various clustering methods, the support vector machines (SVMs), are found to be effective in real-world applications [33, 34]. In addition, the SVMs possess some useful properties for the problems of classification. The simplicity of data driven modeling approach is that there is no need for a fundamental model of the system and only data from normal operation needs to be used, which is generally available in some form for most machines. Among various clustering methods, the support vector machines (SVMs), are found to be effective in real-world applications [33, 34]. In addition, the SVMs possess some useful properties for the problems of classification in terms of the non-linearity, efficiency of computation, and simplicity of implementation [35]. Many applications of this technique have been successfully applied in other fields of process monitoring [36-39]. He and Shi [40] found that support vector machines produced better accuracy than artificial neural networks when applied to a pump diagnosis problem.

This paper outlines the use of the Multi-Class Support Vector Machine SVM approach, to develop a framework to monitor and test the health status of a multi-stage Mechanical transmission system.

II. SYSTEM CONFIGURATION

An automotive mechanical transmissions gear test rig was developed for this ongoing research figure 1. The rig comprises 130 mm centre distance gearbox and fixed on the floor using a non-vibrating platform (Fastened with rubber and bolts).

The system is driven by a 7.5 KW variable speed 3-phase electric motor controlled by an inverter to provide a speed variation of 1750 rpm. The system is loaded through a mechanical braking system and controlled with an AC motor inverter. The system is equipped with five sensors, two accelerometers at two different positions (input and output of the gear system), temperature sensor (immersed in the gearbox oiling system), wireless strain gauge for torque measurements (on the output shaft) and a proximity sensor for speed measurement (at the gearbox input shaft). The rig can generate a load torque on the test gears in the range of 0 – 200 Nm. The torque is measured using calibrated strain gauges installed on the shaft and the measured torque values are transmitted to the control program by telemetry in order to provide torque control of the loading mechanism on the mechanical transmissions. Two temperatures were measured: gearbox oil temperature and bearing temperature using RTD temperature sensors (10mV/C). The input shaft speed and motor current were also monitored as a precaution. The test rig operating conditions were monitored and it is flexibly changed according to the required test conditions using LabVIEW’s virtual instrument scalable architecture features.

![Fig. 1: Multistage gearbox system.](image1)

The Vibration analysis system incorporated a 24-bit NI wireless DSA data acquisition card (NI 9234 with cDAQ-9191) to acquire the vibration signal, speed and temperature. The vibration signals were acquired using two DJB Piezotronic constant current source accelerometers (model no. Acc103 -10mV/g) mounted adjacent to the tested gear bearings transversely to the gearbox casing, and a shaft speed sensor was used to acquire the shaft rotation reference. The sensors location diagram over the test rig is shown in Fig 2. The vibration signals are then acquired continuously and transmitted to the base unit using an IEEE 802.11b/g (Wi-Fi) wireless communication interface (frequency range 2.412–2.462 GHz). The system can send the data from a range up to 30 m for indoor measurements and 100 m for outdoor operation as long as the line of sight of the wireless signal is provided. The system can also provide Ethernet cabling measurements up to a distance of 100 m.

III. SUPPORT VECTOR MACHINES METHODOLOGY

Support vector machines (SVMs) [34, 35] are a group of learning machines for solving pattern recognition problems efficiently. SVMs try to find the hyperplane, which separates optimally the training patterns according to their classes (i.e. hyperplane with maximum boundary margin). This is performed by using what is commonly known in machine learning as the “kernel trick” when using SVM’s. Kernel function is chosen to map the data from its original space to feature space. It can be chosen arbitrarily so as to
best suit the data and at the same time reduce the computational burden involved with generating the mapped values by direct evaluation. “Support vectors” correspond to those points that lie along the margin or closest to it. The maximum margin between classes is found by solving a quadratic optimization problem. SVMs have a good generalization performance over traditional approaches, since their training is based on the principle of structural risk minimization (SRM) (i.e. minimizing the upper bound on the expected risk), while the training traditional approaches are based on empirical risk minimization (i.e. minimizing the number of the training error). SVMs have a high computational efficiency in terms of speed and accuracy.

They are also more preferable when dealing with high dimensional data as they are more robust than traditional approaches which may over-fit the data. However, they still have negative-aspects in terms of giving information about the system output and no physical explanation and interpretation of the process itself. The description of SVMs classification can be explained as follows:

Consider the training data \( \{x_i , y_i\} \), where: \( i = 1, \ldots, N \), \( y_i \in \{+1,-1\} \) corresponding to the class of \( x_i \) (\( y_i = 1 \) for class A, \( y_i = -1 \) for class B). The principle of operation of SVMs classifier will be modified according to the type of the data samples as follows:

**Linearly Separable Data**

Figure 3 shows the hyper plane \( H \) which separates the two classes of data (separating hyper plane). This hyperplane \( H \) satisfies the following equality

\[
b + w^t.x_i = 0
\]

Where: \( w \) is a normal vector on the hyperplane, and \( b \) is a bias representing the distance from the origin.

The training data corresponding to classes A and B satisfy the following inequalities respectively

\[
b + w^t.x_i \geq 1 \quad \text{and} \quad b + w^t.x_i \leq -1
\]

(2)

The two inequalities in (2) can be combined as follows

\[
y_i (b + w^t.x_i) \geq 1
\]

(3)

The two inequalities in (2) can be combined as follows

\[
y_i (b + w^t.x_i) \geq 1
\]

The equalities of (3) define hyperplanes \( H_1 \), and \( H_2 \) respectively, and any training data belongs to class A or class B and lying on \( H_1 \) or \( H_2 \) is called support vectors (SVs). From Fig. 4 the geometry and the separating margin of hyperplane \( H \) is given by

The SVMs classifier tries to find the separating hyperplane with the largest margin (optimal hyperplane). This can be formulated as follows:

**Minimize**

\[
\frac{1}{2}||w||^2
\]

And s.t. constraints in (3)

Using the Lagrangian formulation of the problem

\[
L_p = \frac{1}{2}||w||^2 \cdot \sum_{i=1}^{N} y_i (b + w^t.x_i) + \sum_{i=1}^{N} \alpha_i
\]

(6)

Lp will be minimized with respect to \( w \) and \( b \) and all the derivatives of Lp with respect to all the Lagrangian multipliers, \( \alpha_i \) will vanish. All of these multipliers are subjected to the following constraints:

\[
\alpha_i \geq 1
\]

(7)

The calculations can be simplified by applying Karush-Kuhn-Tucker (KKT) condition which allows applying dual formulation of the problem. This implies that the maximum of Lp is subjected to same constraints in (7) and acquiring that the gradient of Lp with respect to \( w \) and \( b \) vanishes which results in (6) and (7)

\[
w - \sum_{i=1}^{N} y_i \alpha_i x_i = 0
\]

(8)

Using the KKT condition for \( \alpha_i \), we get

\[
\sum_{i=1}^{N} \alpha_i y_i = 0
\]

(9)

Substituting (8) and (9) in (6) results in

\[
L_0 = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \alpha_i (y_i x_i \otimes x_i)
\]

(10)

s.t. constraints in (7) and (9). Once \( \alpha \) is obtained from (10) (using a quadratic programming (QP) solver), the dimensions of the classifier \( w \), \( b \) are determined using (3) and (8). Substituting the obtained values of \( w \) and \( b \) in (11) allows the classification of any unknown sample.

\[
y_{\text{unknown}} = \text{sign} (b + w^t.x_{\text{unknown}})
\]

(11)

The number of variables in (10) is the number of the training data. All the training data associated with the Lagrangian multipliers satisfying the inequality of (7) are the SVs. The number of SVs is considerably less than the number of the training data.

**IV. RESULTS AND DISCUSSION**

This section discusses the results of the experimental
Measurements were driven through three conditions as shown in Fig. 4. shows Kurtosis, Crest factors and temperature of the input gearbox shaft through the different conditions. First, the system was run under normal condition (observations 1-100), then external excitations were applied at one position (input shaft: 101-200). The system was then subjected to a high temperature (201-300).

There are several evaluation schemes for selection of training and testing sets including hold-out, leave-one-out, cross-validation and bootstrap. In our article the selection is based on hold-out method. So, the 300 observations are divided into 225 observations as training set and 75 observations as testing set. The division percentage is 75% for training set to 25% for testing set.

1. Nonlinear support vector machines results

In this section, the generalization of the non-linear SVMs classification algorithm to the gearbox state-of-health data and its performance is investigated. The training technique generated in this design work is adapted to train 225 observations (80 normal, 145 faulty) as training set (Set-1). Subsequently, the model is tested and validated on a subset (Set-T) of the remaining 75 observations (20 normal, 55 faulty) and their corresponding normalized values are directly used as the input features for SVMs. The corresponding output $y_1$ is (1 for a normal condition, 2 and 3 for faulty conditions).

2. SVM models design

The SVM models are designed during the training process by trial and error. The training process involves different Kernel functions as well as several values of each Kernel parameters in order to obtain the SVM classifier with the best performance. The SVM and Kernel methods coded in MATLAB is used for the SVMs training and testing.

3. Selection of Kernel function and Kernel parameters

The SVMs classification technique is tested for two different Kernel functions during the training process namely, the polynomial and Gaussian Radial Basis Functions (RBF) kernel functions. According to the performance of these Kernel functions, the suitability of the SVMs as an intelligent classifier is judged.

The selection of the optimum parameters for SVMs is done during the training process (Set-T). The SVM classifier with the best performance is obtained by testing different values of the Kernel parameters. These parameters are varied in the following manner: $\gamma$ is varied with values of 0.1, 0.2, 0.3, 0.5, 1, 3, 5. The order of the polynomial Kernel $n$ is varied in the range with steps of 2. The penalty due to the error $C$ is also varied with values of 1, 10, 100, 500 and 1000. The tolerance condition for the QP solver is 0.0000001. The performance of the two SVMs is assessed on each of these values by calculating the training percentage performance efficiency defined by:

$$\eta = \frac{\text{No. of samples correctly classified}}{\text{Total No. of samples}} \times 100$$ (12)

From these results, the SVM classifier with the highest training percentage performance efficiency is selected. The testing process is then performed, during which the generalization performance of the classifier is examined using testing set (Set-T) by evaluating the testing percentage performance efficiency.

4. Training and testing the results

Fig. 5 shows the best performance of kernel functions during the training process of the training set (Set-I). The best performance is introduced in terms of the percentage training efficiency of equation (12), with respect to the variation of the kernel parameters $\gamma$, and $n$ as well as the penalty due to the error $C$. The corresponding number of SVs and the training time are also illustrated.
From the results illustrated in Fig. 5 the following clarification are worth noting:

1. Effect of penalty due to the error $C$
   a. For the Kernel functions under investigation, the best performance is obtained at high values of $C = 500$ and 1000. In addition, as $C$ increases the training efficiency increases.
   b. The maximum training efficiency is 96.98% at $C = 1000$.

2. Effect of the Kernel parameters
   a. For the polynomial Kernel, as $n$ increases, both of the number of SVs and the training time decrease, while the training efficiency increases. The best performance for the polynomial Kernel function is 96.13% for $C = 1000$, $n = 10$.
   b. For the Gaussian Kernel, as $\gamma$ decreases, both of the number of SVs and the training efficiency increase, while the training time decreases. The best performance for the Gaussian Kernel function is 96.98% for $C = 1000$, $\gamma = 0.1$.

3. Effect of the type of the kernel function
   a. The best training efficiency was obtained with the Gaussian Kernel function (96.98% during SVM training).
   b. The shortest training time was obtained for the polynomial Kernel function (7.4 second during SVM training and the smallest number of SVs was obtained for both the polynomial and Gaussian Kernel function (10 SVs during SVM training).

Figures 6 and 7 demonstrate samples of contour plots for non-linear SVM classifier using Gaussian Kernel functions. $C=1000$, Gaussian, $KO = 2$, lambda=$1e-7$
$C=1000$, Polynomial, $KO = 2$, lambda=$1e-7$

V. CONCLUSIONS

The study has presented a wireless vibration measuring system that was able to detect different conditions of gears in automotive gearbox and clearly classify its condition using one accelerometers at input of the gear system, temperature sensor (immersed in the gearbox oiling system), wireless strain gauge for torque measurements (on the output shaft) and a proximity sensor for speed measurement (at the gearbox input shaft) for model building and testing. The study has focused on monitoring the classifying of system faults using multiclass SVM.

The system is being developed for use on 130mm automotive manual transmissions, but could be adapted for other transmission or machinery systems rotating machinery.

The model was tested under different conditions including: normal condition, external vibrational excitations at one position (input shaft), high temperature and was able to successfully differentiate between them.

VI. REFERENCES


