APPLICATION OF MACHINE LEARNING FOR MACHINE MONITORING AND DIAGNOSIS

David A. Maluf
Stanford University, Stanford
maluf@db.stanford.edu

Laeeque K. Daneshmend
Queens University, Canada
lkd1@post.queensu.ca

Abstract

This paper reports on research to evaluate the application of artificial neural networks to pump condition monitoring. Based on historical velocity vibration measurements, artificial neural networks were developed to assess pump condition. Pumps, in general, behave like most rotating equipment and evolve or deteriorate as a function of a single variable, namely time. Such evolution forms the basis of the current work, where the monitoring of the vibration measurements at different times establishes a pattern in that evolution. Therefore, given adequate data of the initial life, optimum behavior, and final stages of machine failure of the pump, the relative pump health can be determined.

1 INTRODUCTION

The focus of this paper relates to the research effort to examine historical vibration data monitored from rotary equipment to determine the feasibility of applying neural networks to identify machine condition. This paper outlines some results of this effort in which artificial neural networks (ANNs) were used to establish pump condition at the Montreal Shell Refinery and were based on frequency characteristics of monitored vibration data. In addition, the emphasis of this case study was on the Bingham pump types chosen arbitrarily from a wide variety of pumps being monitored at the Shell facility. This particular type of pump was selected for study as there was a comprehensive set of vibration data available covering an historical period of one year. An on-site monitoring system which has been used by Shell primarily for computer aided spectrum comparison and prediction, has provided for the current study the minimum required historical data. The data was acquired using the Computational Systems Incorporated (CSI) products and environment [3], which provide range of vibration velocities from the frequency spectrum. In addition to this vibration data, Shell also provided detailed maintenance logs providing information on the pump type, date and type of failure. The pump monitoring approach outlined in this paper is distinct from previous efforts in that it involves the classification of the velocities into patterns through artificial neural network systems which can characterize the condition of the pump that is being monitored.

The process outlined in this research starts with a thorough inspection of the site measurements cross-referenced to the maintenance logs. Faults describing *unbalance*, *mis-alignment*, *gearbox* failure and *bearing* related damages were tagged and the corresponding historical measurements were selected. "MasterTrend", a vibration analysis software package designed to be operated under Computational Systems Inc. environment, was used to establish the front-end interface to the vibration measurements [3].

Research involving the use of artificial neural networks requires an extensive effort to design an efficient and functional artificial neural network [2,5]. Variables such as number of neurons and the training methods require in-depth experimentation in order to achieve optimal results. Specifically, this paper describes an approach to deal with problems in using neural network in studies of sparse data. Although many different tests were carried out for comparison and evaluation purposes, this paper will describe only a typical example in using neural networks for fault detection, prediction and diagnosis.

2 DATA ACQUISITION

Vibration monitoring can be accomplished through the measurements of the parameters that govern the dynamics of the pump [6] as shown in Figure 1. In fact, the selection among the displacement, velocity and acceleration measurements is usually arbitrary. Nevertheless, from a practical point of view, each kind of measurement exhibits a different amplitude response to frequency.

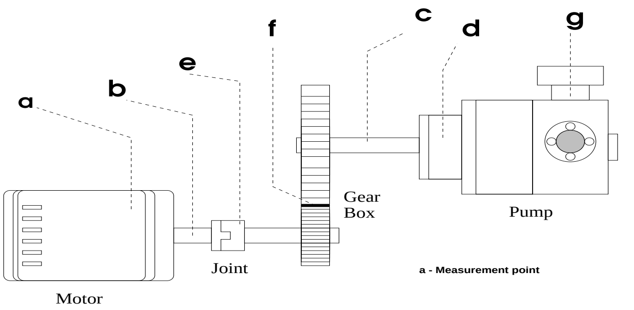

Figure 1: Vibration is measured at different point selected around the the overall pump setup which consists of a motor, joint, gear box and a pump.

In general, the velocity measurements provide a wideband, reliable frequency spectrum compared to displacement and acceleration measurements which are reliable in low and high frequency spectra respectively [1]. There are in fact direct mathematical relationships between the parameters. However, in practice, it is preferable to directly acquire the raw data in the preferred format.

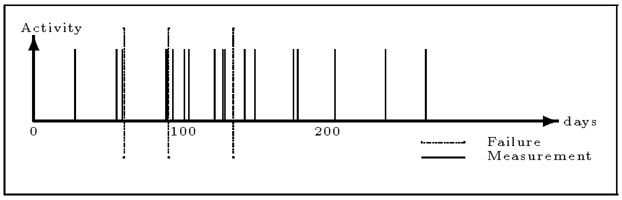

Figure 2: Time delay between successive measurements (pump x2304).

The selection of pumps as the focus of this study and in particular the selection of the Bingham pump was largely arbitrary. The measurements of the Bingham pump labeled "x2304" in Shell East End Montreal Refinery logs, were selected as the main resource for this paper. Hence, the monitoring logs for the "x2304" pump was extracted from the database. Figure 2 shows the time delays between consecutive measurements on that pump, where the delay is a function of days. The numbers of measurements are too low to justify a statistical analysis, and the time delay between the nineteen measurements varies from a minimum of less than an hour to thirty five days. In fact, for the selected pump the relationship between the distribution of failures and the frequency of measurements are correlated, since the maintenance personnel realized some kind of malfunction was imminent and measured the vibration more often. Figure 2 shows that the faults occured within a few days of the most recent measurements. whereas elsewhere the monitoring was quite sparse.

The monitoring and data acquisition procedure followed to obtain the data can be summarized as follows. The pumps which are to be monitored are selected and grouped into sites. Each pump is then considered individually and points on it are permanently chosen in the expectation of useful measurements as in Figure 1. Such points are chosen based on manufacturer recommendations. Since vibrations are the outcome of the force transmission through the pump, the direction of the vibration is unknown. Hence, measurements were taken in the three orthogonal directions (horizontal, vertical and axial). Each measurement consists of 398 pairs of a frequency and the corresponding root mean square velocity vibration which was measured in inches per second. The overall bandwidth varies from 300 Cycles Per Minute (CPM)(5 Hz) to 150,000 CPM (2500 Hz), with an increment resolution approximately of 376 CPM (6 Hz).

3 PROCESSING METHODOLOGY

The context of evaluating the application of neural network to pump monitoring can be directly related to the methods involved in the integration of the measurements to the neural network [4]. Dimensional mapping and similar procedures are to be handled carefully according to the requirements of each of the neural network design. However, to determine a good neural network design, and specially in the current case where the number of samples is small, a strategy has to be devised in partitioning the data measurement.

3.1 Classification

It has been established that the health state of pumps is characterized by unique vibration patterns which are often called signatures. Tracking these signatures can provide an indication of the pumps health. Consistency in the measurements is necessary for the neural network to converge and thus classify the patterns. In other words, contradictory training brings the neural network to a long and exhaustive learning phase and eventually to a convergence failure. On the other hand, consistent measurements generate a robust neural network and quick adaptation.

The strategy in organizing the data provides a potential effect on the final validation of the overall neural network performance. The sensitivity of the neural network to data grouping is due to the small amount of pump failures. In the current case study with Shell, failures constituted less than 5% of all the measurements. In this situation where the total amount of measurements were scarce, the grouping of failures to less complex categories was proposed. The following classifications were attempted. (i) Classification of two classes: *healthy*, *non-healthy* pump. (ii) Classification of four classes: *unbalance*, *mis-alignment*, *gearbox*, *bearing*.

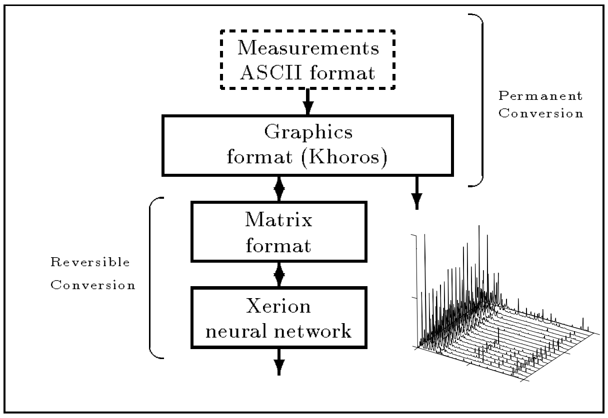

Figure 3: Software architecture. Khoros and Xerion are public domain software. Xerion was developed by G. Hinton at the University of Toronto. Khoros is best described in [7].

3.2 Diagnosis prediction and labeling

In this paper we consider the measurements taken at the instance prior to a failure are the *absolute* description of the machine before the occurrence of a fault. The concept of absolute relate to the fact that the time of last measurement and the occurrence failure is minimal. On the other, we consider all other measurements prior to a final measurement are *relative* as a function of the time period between the measurements and the fault. In a practical manner, the existence of absolute measurements are only feasible through continuous monitoring. Relative measurements are obtained by more flexible methods varying from continuous to routine procedures. The measurements used in the current work are considered relative and weak, in general due to the effect of the elapsed time between the measurements and the fault. However, since the intention of the current work is to assess the pump health, this objective is close enough to the concept of generating a comparative scale as a function of the relative measurements.

Relative assessment can be denoted as percentage for each of proposed classes. Here, the assumption of setting the value "1" (100%) as the worst health or failure and "0" (0%) as the best health are defined. It becomes easy enough to characterize the health of the pump operating between the two mentioned boundaries, namely an index the machine life cycle. For example, given 75% unbalance an operator can assume that the machine has spent 75% of its life cycle. The operator may correlate an approximate time to schedule its maintenance.

3.3 Understanding the data

The learning algorithm for the training of a neural network used in the current work was based on the *back propagation* of the error. A single input vector (data) is presented to a neural network, the output vector is calculated and the training algorithm determines the weight updates of the network. However, different techniques in handling the training algorithm can lead to higher efficiency, and hence, multiple inputs can be presented simultaneously. This is known as *batching*.

Taking advantage of the flexibility of the environment used for the current simulation, a visual assessment was carried out throughout the simulation process as shown in Figure 3. *Visual assessment* is a primitive way to display the data in the frequency spectrum where peaks are identified. In practice, visual assessment is a dominant approach in machine diagnostics which are enhanced with thresholding labeling methods combined with alarms [3]. Visual assessment has much to do with computer graphics and displays [7]. Once the measurements are encoded in standard format for the analysis, visual assessment can be achieved simultaneously with the neural network diagnostic.

4 SOME RESULTS

Considering the fact that vibration signature patterns vary within small ranges, and that they also reflect the signals energy, averaging techniques and dimension mapping are reasonably justified. However the question can be raised about the benefit of similar techniques. The investigation described in the current section involves a comparative test for different neural networks designed for different input vector dimensions. The comparison is perceived from the training and testing point of views.

The different tests were undertaken on set percentages of the complete dimensions, that is, considering the initial vector dimension of 398 points (as presented earlier) we reduced them to a final stage of forty points with constant reduction intervals of forty (10 percent of the initial value). The dimension mapping was completed based on octave bands using both Hamming and square windows. There was ten windows in total. We consider windowing as an averaging technique with partial overlaps. In this paper, the overlap was taken to be fixed at 30 percent of the window size.

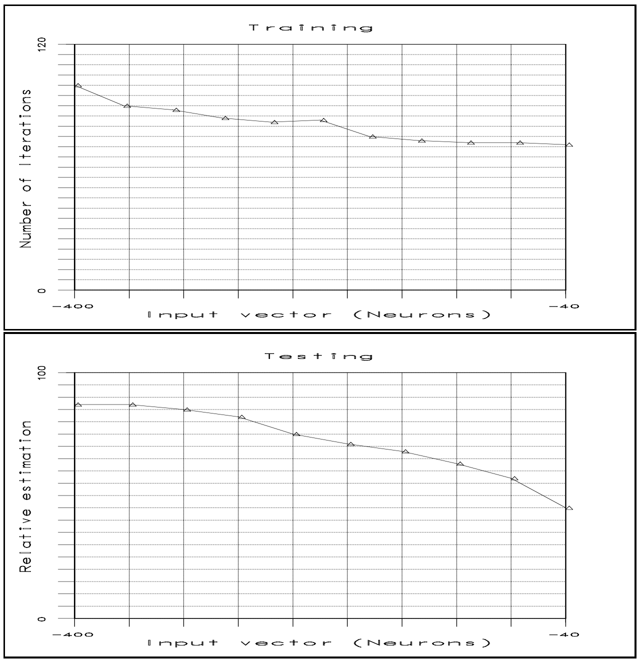

Figure 4: Tracking the effect on vector mapping in both training and testing.

The test that was achieved in the current section was based solely on the Bingham pump "x2304". The mentioned pump suffered a shaft failure and was closely monitored thus providing adequate description of the failure. For the proposed fault, a neural network output layer was designed to handle three outputs. The first output neuron was trained to exhibit a main alarm varying between two values of zero and one. Similarly, the second output neuron was trained to behave also in a Boolean fashion and to show that the fault is a shaft failure. Finally, the last output neuron was trained to exhibit the relative shaft health in a percentage fashion. As opposed to the type of classes proposed in section 3.2, the proposed output neurons was selected to present a subset of them.

The data measurements were arbitrary divided into two groups. The first group consisted of 74% of all the measurements and was employed for the training session, whereas the testing session was based on the rest of the measurements. One should realize that since the data were sub-sampled, the temporal effect was lost and thus testing with recurrent neural networks was not valid for the current simulation.

Figure 4 shows the result of the simulation. The first plot shows the variation in the amount of training iterations versus a decrease in the input vector dimension. For the ten tests that were carried out, an approximate drop of 30% in the number of iterations was perceived. Such results were foreseen since the signature patterns become simpler and thus easier to classify.

For each time a new neural network was designed and successfully trained, the selected network was tested with the other 26% of samples of all the measurements. The bottom plot in Figure 4 exhibits the classification performance of the second output neuron once faulty measurements were presented to the input layer. The network was trained to exhibit values larger than ninety. Once the simulation was completed, the second neuron of the output layer was recovered and plotted as a function of the dimension of the input vector. For a full dimension (398 point input vector), the neural network has detected the failure with an accuracy of 87%. However, for a forty point input vector, the neural network has detected the failure with an accuracy of 54%. The drop between these two extremes was steady leading to the conclusion that the decrease in the dimension resulted in weaker pattern classifications.

The results in Figure 4 prove that a decrease in the input vector dimensions results in faster training in the context of the number of iterations required. Moreover, it is known that the computational complexity for each neural network design in the context of floating point calculations is linearly dependent on the dimension of the input vector. In other words, reducing the input vector from 398 to forty decreases the network computational complexity by at least ten times (given the same output requirements). On the other hand, the proper evaluation of the neural network performance is highly dependent on the output behavior. Although the above experiment indicates a requirement of larger input dimensions, large input vector require extensive amount of data measurements to train with.

4.1 The *Weight ties* approach

This section shows a heuristic method used to reduced the effective input matrix dimension. Two data sets for training and testing were used in these tests with the intention of comparing the results. The innovation introduced in the current section is based on grouping the measurement values mapped to the same frequency from different measurement points (see Figure 1). Forty *ties* were used to group the averaged measurements from the same octave frequency spectrum. Although the number of measurement points varies from a machine to another, for the described experiment a total number of seven measurement points given three directions were considered for a total of 18 vectors. The considered directions were horizontal, vertical and axial. Three axial measurements were not feasible. However, it was left to the training session to determine the values carried by the mentioned weights.

Pump		*Network*	*Classification*	
Training	*Testing*	*architecture*	*Fault*	*Error*
x2304	x2304	Feed-forward	Shaft	07%
x2304	x2304	Feed-forward	Joint	12%

Table 1: Simulation results based on weight ties method.

In Table 1, the training samples consisted of 74% of the data measurements as mentioned earlier.Compared with Figure 4, it is clear that the perspective of cross measurement training increases the classification quality. As shown in Table 1, the weight tie method preserved the cross measurement relations, hence increasing the neural network performance.

5 CONCLUSION

One of the advantages of neural network applications is that their performance has a tendency to increase with larger numbers of measurements, assuming that the training time period is not constrained. In addition, current computer technology has enormously improved the processing time involved in training.

To enhance the classification of the neural network it is preferable to present to the neural network measurements of faulty pumps just before failure and the measurements after the pumps are repaired. Such an approach is generally difficult to implement. However, to compensate to the difficulty, domain experts can adequate label the training of the neural network and hence minimize the problem. In fact it is very difficult to identify the measurements that describe a complete failure, rather than those that describe the relative health. As a result it is necessary to collect measurements as close in time for a failure to occur for properly labeling the data while training the neural network. To approach the described limit can be achieved by decreasing the time between consecutive measurements.

A safe margin for separation time can be estimated and defined to ensure that the necessary amount of data are collected for the neural network to exhibit satisfactory results. Also from historical facts, the minimum period of time between any two failures or MTTF (Minimum Time To Failure) can be computed where the sampling period of measurement is less than half of the computed period. In other words, given a certain pump, with a minimum elapsed time between any two failures, measurements should be taken at least twice in that period of time. In addition, all monitoring data should be collected over adequately spaced intervals for the selected period of time. The sampling period would have to be significantly reduced for preventive maintenance purposes. In addition, given the nature of the current application, we believe data records should number in the hundreds in order to train dependable neural networks.

Eventually, continuous monitoring could become the most reliable and accurate method of pump health monitoring. On-line failure detection minimize the shut down delay time which might be crucial in some cases for hardware salvage and safety. In the case where continuous pump monitoring is difficult to implement, concessions become obvious through the use of off-line approaches. Neural networks are flexible and can be adapted to numerous applications. However, fundamental requirements are needed to guarantee the robustness of the decision making entrusted to the artificial neural networks.

References

[1] Bruël & Kjaer. Machine-Health Monitoring, Application note BR0267-12.

[2] Benjio Y., Artificial Neural Networks and their Application to sequence Recognition, *Ph.D dissertation. Dept. of Comp. Sci., Mcgill University*, 1991.

[3] Computational Systems Incorporated, *Master Trend*, Software Package for Predictive Maintenance, Vol1 & Vol2, 1992.

[4] Cumming, S., Neural Networks for Monitoring of Engine Condition Data, *N. Comp. & App.*, vol. 2, pp. 96-102, 1993.

[5] Gallant, S.I., Neural Network Learning and Expert Systems, *MIT Press, ISBN 0-262-07145-2*, 1993.

[6] Little, P. D. and Singh, R. Turbine Engine Degradation on Life Usage, *Proc. of the Eighth Int. Cong. on Condition Monitoring and Diagnostic Engineering Management*, vol. 2, pp. 353-360, June 1995.

[7] J. Rasure, D. Argiro, T. Sauer, and C. Williams. A Visual Language and Software Development Environment for Image Processing, *International Journal of Imaging Systems and Technology*, Vol. 2, pp 183-199, 1990.