A Neural Network-based Application for Oil and Gas Pipeline Defect Depth Estimation

Abduljalil Mohamed*, Mohamed Salah Hamdi*, and Sofiene Tahar**

* Ahmed Bin Mohamed Military College/Information Systems Department, Doha, Qatar
** Concordia University/Electrical and Computer Engineering Department, Montreal, Canada
{ajmaoham, mshamdi}@abmmc.edu.qa
tahar@ece.concordia.ca

Abstract—Experienced engineers utilize Magnetic Flux Leakage (MFL) sensors to scan oil and gas pipelines for the purpose of localizing and sizing different defect types. The huge amount of raw data obtained by these sensors, however, makes the inspection task error-prone and time-consuming. In this paper, we propose a defect depth estimation approach using artificial neural networks of various architectures. Discriminant features, which characterize different defect depth patterns, are first obtained from the raw data. Neural networks are then trained using these features. The Levenberg-Marquardt back-propagation learning algorithm is adopted in the training process, during which the weight and bias parameters of the networks are tuned to optimize their performances. Compared with the performance of pipeline inspection techniques reported by service providers such as GE and ROSEN, the results obtained using the method we proposed are promising. For instance, within ±10% error-tolerance range, the obtained estimation accuracy is 86%, compared to only 80% reported by GE; and within ±15% error-tolerance range, the achieved estimation accuracy is 89% compared to 80% reported by ROSEN.

I. INTRODUCTION

Natural gas and crude oil production is usually carried through long-distance transmission metallic pipelines. Due to the nature of environment and extreme temperature, metallic pipelines are subjected to corrosion, which is considered as a leading cause of pipeline defects. In [1], it has been reported that nearly 30% of pipeline defects are due to external corrosion of underground or underwater pipelines. These pipeline defects can result in huge financial losses, damage to the environment, and loss of life. Thus, pipeline operators are required to utilize effective and efficient intelligent tools to detect and locate pipeline defects. Efficient intelligent tools utilize Magnetic Flux Leakage (MFL) signals and ultrasonic waves and use them to detect and localize defect types (e.g., corrosion, cracks, dents, etc.). MFL recordings around the center of a metal-loss defect do have a distinct pattern of behavior. The sensor passing directly above the defect center has highest amplitude of the axial and radial components of the MFL signal. The amplitude of these components decrease further away from the defect center. Using the MFL measurements of the neighborhood sensors, the type and size of a defect can be determined. In the literature, several techniques have been proposed for the purpose of detecting and localizing pipeline defects. In [2], using MFL signals, Artificial Neural Networks (ANNs) are used to classify signal patterns with three types of defects in the weld joint of pipelines: External Corrosion (EC), Internal Corrosion (IC) and Lack of Penetration (LP). The defects are intentionally inserted in the weld bead of a pipeline. The results have showed that it is possible to classify signals with non-defect (ND) and signals with defects (D) along of the weld bead using ANNs with very high accuracy (94.2%). For corrosion (CO) and LP signals the accuracy was 92.5%. In [3] a fuzzy artificial neural network–based approach is proposed for reliability assessment of oil and gas pipelines. The actual condition of aging pipelines vulnerable to metal loss corrosion are characterized by eight pipe parameters as input variables obtained from MFL signals. The proposed method uses these parameters to estimate the probability of failure of aging pipelines vulnerable to corrosion. In [4] a recognition and classification of pipe cracks using images analysis and neuro-fuzzy algorithm is proposed. Crack-related features are first extracted. The combination of a fuzzy membership function, used to absorb variation of feature values, and a back-propagation network, with learning ability, shows good classification efficiency. In [5], a radial basis function neural network (RBFN) is deemed to be a suitable technique and a corrosion inspection tool to recognize and quantify the corrosion characteristics. An Immune RBFNN (IRBFNN) algorithm is proposed to process the MFL data to determine the location and size of the corrosion spots on the pipeline.

Moreover, there is a relationship between the amplitude and the area under the curve of the MFL signal and the depth of the corresponding defect. Since this relationship is not well-understood and cannot be analytically described, neural networks can be trained to learn this relationship to estimate a defect depth. To the best of our knowledge, neural networks have only been used to detect and localize pipelines defects. In this paper, the use of artificial neural networks (ANNs) for estimating pipeline defect depths is evaluated, using magnetic flux leakage (MFL) signals obtained by an intelligent pig.
II. PIPELINE DEFECT DETECTION AND SIZING

A. MFL-based Pipeline Inspection

Nondestructive pipeline inspection uses autonomous devices equipped with strong magnets and magnetic sensors, called intelligent pigs, which are sent into and retrieved from pipes to measure any magnetic flux leakage. The sensors are equally-distributed around the circumference of the pipeline, and move with the intelligent pig parallel to the axis of the pipeline. Intelligent pigs utilize MFL scanning. The main concept behind MFL scanning is the following [6]. When two strong magnets of opposite polarity are held close to the surface of a pipeline, the latter gets magnetized, and lines of magnetic force (called magnetic flux) flow through the walls of the pipeline, from the south pole to the north pole. When the pipeline wall contains a crack or a thinning (due to corrosion, for example), then at the edges of the crack two new poles appear. However, the air gap between the two new edges cannot absorb as much flux per unit of volume as the ferromagnetic material and, hence, causes the magnetic lines of force to bulge out. The bulging of magnetic flux is called Magnetic Flux Leakage (MFL). Any magnetic flux leakage detected by the sensors indicates the presence of a defect. The MFL signals measured by the sensors are recorded and later analyzed to locate possible defects, and determine their sizes and severity levels.

B. Metal-loss Detection and Sizing using Wavelet-based Techniques

Wavelets [13, 14, 15] are a powerful mathematical tool with a wide variety of applications ranging from high-efficiency data compression [16] to data analysis and classification [17]. They are also widely used as an efficient de-noising technique [7, 8, 9] and sometimes, they are combined with other tools and used for detecting the presence of metal-loss defects [10]. In our work, we use pattern-adapted wavelets for locating metal-loss defects and estimating their length. MFL signals take a certain shape at the location of a metal-loss defect. Furthermore, this same shape occurs in a dilated form depending on the length of the defect. Let \( B(x) \) denote the MFL signal measured from a pipeline. Figure 1 shows a sample MFL scan containing three defects of cuboidal shape. Each of these defects has a different length along the x-axis of the pipeline. All three components of the MFL signal \( (B_x, B_y, B_z) \) are represented. As can be seen in Figure 1, each component of the MFL signal consists of a sum of curves, each of which is translated and dilated version of a reference pattern. If we choose the reference pattern as a mother wavelet, \( \psi(x) \) and derive an orthonormal wavelet basis \( \langle \psi_{j,k}(x) \rangle \) from it, then the MFL scan \( B(x) \) can be expressed in the basis as:

\[
B(x) = \sum_{j,k} c_{j,k} \psi_{j,k}(x).
\]

The non-zero coefficients \( c_{j,k} \) in the above representation indicate that the signal contains a copy of that particular instance of \( \psi_{j,k}(x) \). To detect a metal-loss defect along the pipeline, and estimate their length, the wavelet transform of the MFL scan \( B(x) \) is first computed with respect to the basis \( \langle \psi_{j,k}(x) \rangle \). Next, the set of non-zero coefficients, which indicates the locations and dilation factors of the reference pattern, is determined. This in turn yields the location and width of all metal-loss defect along the pipeline.

III. NEURAL NETWORK-BASED APPROACH FOR ESTIMATING DEFECT DEPTHS

Along with the length of the defect, its depth is a very important factor for determining its severity. According to industry standards, some defects, based on their depths, may be considered completely safe, while others are deemed too severe. It has been observed that the magnitude of MFL signals is much higher for defects with larger depths. The relationship, however, between defect depths and the magnitude of MFL signals is not well understood and cannot be analytically described. Therefore, machine learning techniques can be used to capture this relationship. In this paper, we investigate the application of Artificial Neural Networks (ANNs) as a learning tool and propose an ANN-based approach for estimating failure depths. The structure diagram of the proposed approach is shown in Figure 2.

As shown in Figure 2, a feature set is first extracted from the MFL signals. These features should be representative of patterns of different failure depths. These features are then used in the learning process of the neural network.

A. Feature Extraction

Feeding raw MFL signals directly into the neural network may prolong the learning task and lead to unsatisfactory results. Instead, a number of features that characterize the MFL signals are first calculated. In this work, the following features are extracted:

- Maximum magnitude
- Peak-to-peak distance
- Integral of the normalized signal
- Mean average
- Standard deviation
Moreover, polynomial series of the form, 

\[ a_nX^n + \ldots + a_1X + a_0 \]

are used to approximate the MFL signals. In particular, polynomials of degrees 3, 6, and 6 have been found to provide the best approximation for \( B_x, B_y \) and \( B_z \) respectively. The polynomial coefficients \( a_n + \ldots + a_0 \) have been considered as features and used, along with the above features, to train the neural network. In total, we have an input vector consisting of thirty-three features.

B. Neural Network Architecture and Parameters

We examine three architectures, namely static, cascaded, and dynamic feed-forward neural networks (FFNN).

1) Static FFNN
The architecture of the static FFNN is shown in Figure 3. The extracted feature vector is fed into the first hidden layer. Weight connections, based on the number of neurons in each layer, are assigned between every adjacent layers.

2) Cascaded FFNN
As shown in Figure 4, these networks are similar to feed-forward networks, but include a weight connection from the input layer to each other layer, and from each layer to the successive layers.

3) Dynamic FFNN
In dynamic networks as shown in Figure 5, the network outputs depend not only on the current input feature vector, but also on the previous inputs and outputs of the network.

4) Neural Network Parameters
In this study, different network parameters are examined including the number and size of hidden layers, the type of the transfer functions, and the type of the performance functions. As for the opted learning algorithm, it has been shown that the Levenberg-Marquardt back-propagation algorithm provides the best performance for function approximation; and hence it is more suitable than other learning algorithms for defect depth estimation [11, 12].

IV. PERFORMANCE EVALUATION

The main performance measure used to evaluate a given network structure and configuration is the estimation accuracy of the failure depth within a certain level of error-tolerance. The error-tolerance levels used in this study are \( \pm 1\%, \pm 5\%, \pm 10\%, \pm 15\%, \pm 20\%, \pm 25\%, \pm 30\%, \pm 35\%, \pm 40\% \). For each network structure, the FFNN is experimented with different numbers of hidden layers, each varies in size from 10 neurons up to 100 neurons. The results of the experimental work are reported in the following subsections.

A. Performance and Transfer Functions

The performance function is the first parameter examined as it plays a crucial role in the accuracy and speed of network learning. Two performance functions are selected:

1) Mean Squared Error (MSE)

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2
\]

2) Sum Squared Error (SSE)

\[
SSE = \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2
\]

The effect of the MSE performance function is first examined on a static FFNN with a single hidden layer and different number of neurons. The network performance is tested on two types of transfer functions used in neurons of the hidden layer:

1) Log sigmoid

\[
\log \text{sig} = \frac{1}{(1-e^{-x})}
\]

2) Hyperbolic Tangent sigmoid

\[
\log \text{sig} = \frac{2}{(1+e^{-x})} - 1
\]

Figure 6 shows the results obtained by the static FFNN with MSE as a performance function and Log-sigmoid as a transfer function. It can be noted from the figure that, for \( \pm 1\% \) error-tolerance level, the defect depth estimation accuracy obtained by the network is approximately 10.29% with 50 neurons in the hidden layer. Naturally, as the error-tolerance increases the network performance starts getting better. For instance, for \( \pm 5\% \) error-tolerance level, the defect depth estimation accuracy obtained by the network, with 70 neurons in the hidden layer, is around 47.55%. For \( \pm 10\%, \pm 15\%, \pm 20\%, \pm 25\%, \pm 30\%, \pm 35\%, \pm 40\% \) levels, the network best estimation performances are 64.71%, 79.41%, 83.33%, 88.24%, 90.69%, 90.69%, 91.18% with 60, 20, 20, 10, 10, 80, and 80 neurons in the hidden layer, respectively.

Figure 7 shows the results obtained by the static FFNN with SSE as a performance function and Log-sigmoid as a transfer function. For \( \pm 1\%, \pm 5\%, \pm 10\%, \pm 15\%, \pm 20\%, \pm 25\%, \pm 30\%, \pm 35\%, \pm 40\% \) levels, the network best estimation performances are 64.71%, 79.41%, 83.33%, 88.24%, 90.69%, 90.69%, 91.18% with 10, 20, 20, 10, 10, 80, and 80 neurons in the hidden layer, respectively.

Tables I and II show the comparison of results obtained by the static FFNN using different transfer and performance functions. For both performance functions (MSE and SSE), both transfer functions show very close performance results. However, the difference in hidden neurons is notable. For example, at \( \pm 1\% \) level error-
tolerance, an FFNN with Log-sig transfer and MSE performance functions needs only 10 neurons to obtain accuracy of 10.29%, while with Tan-sig transfer function, an FFNN needs 70 neurons to obtain 9.80% accuracy. At ±5% error-tolerance level, the network configuration is the opposite, an FFNN with Log-sig transfer and MSE performance functions needs 70 neurons to obtain 47.55% accuracy, while with Tan-sig transfer function, an FFNN needs 70 neurons to obtain 45.59% accuracy. Similar observation can be made for the FFNN with SSE performance function.

B. Static, Cascaded, and Dynamic FFNN

Based on the results and observations obtained in the previous section, the MSE performance function and Log-sigmoid transfer function are fixed for the next experiment settings. Table III shows the comparison of the results obtained by the static, cascaded, and dynamic FFNN structures with different numbers of hidden layers.

It should be noted from table III that dynamic networks with a single hidden layer yield the best performance results for error-tolerance levels of ±1%, ±5%, ±10%, ±15%, and ±20% at 23%, 74%, 86%, 89%, and 90% estimation accuracies, respectively. Moreover, dynamic networks with 4 hidden layers yield the best performance for error-tolerance levels of ±25%, ±30%, ±35%, and ±40%, at 91%, 93%, 95%, 96%, and 96% estimation accuracies, respectively. Cascaded networks, however, have performed the worst for error-tolerance levels of ±1%, ±5%, ±10%, ±15%, and ±20%, at 7%, 4%, 60%, 72%, and 78% estimation accuracies, respectively. At other error-tolerance levels, they yield comparable results. Static networks performed better than cascaded networks but less than dynamic networks.

It should be noted that, in this particular application, increasing the number of hidden layers has not necessarily improved the performances of the networks.
TABLE III.
BEST ESTIMATION ACCURACY OF STATIC, CASCADED, DYNAMIC FFNN

<table>
<thead>
<tr>
<th>Error Tolerance</th>
<th>Static</th>
<th>Cascaded</th>
<th>Dynamic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hidden Layers</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>0.1%</td>
<td>0.81</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>0.5%</td>
<td>0.81</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>1%</td>
<td>0.81</td>
<td>0.81</td>
<td>0.81</td>
</tr>
</tbody>
</table>

With the exception of dynamic networks (with 4 hidden layers, and for the error-tolerance levels ±1% and ±5%), it has actually reduced the overall performance of the feed-forward neural networks.

C. Discussion

We suggested an approach consisting of combining pattern-adapted wavelets for locating metal-loss defects and determining their length and neural networks for estimating the defect depth. In the current work the focus was on studying the applicability and suitability of the neural network component. Compared with the performance of pipeline inspection techniques reported by service providers such as GE and ROSEN, the results obtained using the method we proposed are promising. For instance, within ±10% error-tolerance range, the obtained estimation accuracy is 86%, compared to only 80% reported by GE; and within ±15% error-tolerance range, the achieved estimation accuracy is 89% compared to 80% reported by ROSEN.

V. CONCLUSIONS

The use of artificial neural networks (ANNs) for estimating pipeline defect depths is evaluated using different levels of error-tolerance. Extensive experimental work for different parameters and configurations, and various architectures including static, cascaded, and dynamic networks, has been conducted. The Levenberg-Marquardt back-propagation learning algorithm has been utilized. It has been shown that dynamic neural networks yield the best performance of 86% and 89% defect depth estimation accuracy within ±10% and ±15% of error-tolerance, respectively; while cascaded networks yield the worst performance. To increase the defect depth estimation accuracy at lower levels of error-tolerance, we intend in future work to obtain more sophisticated features and employ other learning algorithms.

ACKNOWLEDGMENT

This work was made possible by NPRP grant # [5-813-1-134] from Qatar Research Fund (a member of Qatar Foundation). The findings achieved herein are solely the responsibility of the authors.

REFERENCES


