Ultrasonics 54 (2014) 1534–1544

Contents lists available at ScienceDirect

Ultrasonics

journal homepage: www.elsevier.com/locate/ultras

Active incremental Support Vector Machine for oil and gas pipeline defects prediction system using long range ultrasonic transducers

Nik Ahmad Akram [a], Dino Isa [a], Rajprasad Rajkumar [a], Lam Hong Lee [b,*]

[a] The University of Nottingham Malaysia Campus, Jalan Broga, 43500 Semenyih, Selangor Darul Ehsan, Malaysia

[b] Quest International University Perak, No. 227, Plaza Teh Teng Seng, Level 2, Jalan Raja Permaisuri Bainun, 30250 Ipoh, Perak, Malaysia

ARTICLE INFO

Article history:
Received 13 January 2014
Received in revised form 10 March 2014
Accepted 31 March 2014
Available online 18 April 2014

Keywords:
Support Vector Machine
Oil and gas pipeline
Non-Destructive Testing
Artificial intelligence
Incremental learning

ABSTRACT

This work proposes a long range ultrasonic transducers technique in conjunction with an active incremental Support Vector Machine (SVM) classification approach that is used for real-time pipeline defects prediction and condition monitoring. Oil and gas pipeline defects are detected using various techniques. One of the most prevalent techniques is the use of "smart pigs" to travel along the pipeline and detect defects using various types of sensors such as magnetic sensors and eddy-current sensors. A critical short coming of "smart pigs" is the inability to monitor continuously and predict the onset of defects. The emergence of permanently installed long range ultrasonics transducers systems enable continuous monitoring to be achieved. The needs for and the challenges of the proposed technique are presented. The experimental results show that the proposed technique achieves comparable classification accuracy as when batch training is used, while the computational time is decreased, using 56 feature data points acquired from a lab-scale pipeline defect generating experimental rig.

1. Introduction

This work presents an oil and gas pipeline defect prediction system using Long Range Ultrasonic Testing (LRUT) in conjunction with an active incremental Support Vector Machine (SVM) technique. Pipeline systems are normally safe and more effective method for mass-scale delivery of gas and liquid formed products. According to USA Association of Oil Pipelines, the system loses around 1 gallon per million barrel-miles (One barrel-mile = one barrel transported a mile) where a standard barrel contains 42 gallons (159 l) [1]. In other words, less than one teaspoon of oil is spilled per thousand barrel-miles. Pipeline systems are considered as the most cost-effective operation as compared to railway and road transportation in the long term [2].

The main problem faced by the oil and gas industry, having been in operation for 185 years, is that aging pipeline systems are being corroded and are undergoing wall thinning which can eventually lead to pipeline failure [3]. The main causes of pipeline failures around the world are corrosion defects such as cracking, pitting and Stress Corrosion Cracking (SCC).

Some of the causes of pipeline failures can be avoided especially corrosion [4]. Table 1 shows that corrosion causes 18.4% of total accidents. Currently, pipeline inspection is done at predetermined intervals using techniques such as pigging where operators must be physically present to perform measurements and make judgments on the integrity of the pipes. The condition of the pipe between these testing periods, which can be for several months, can go unmonitored. The number of factors that cause failures in pipes are usually large and mostly unexpected. This is the main cause for frequent pipe failures and leaks, as the defects which lead to failure usually occur suddenly. Therefore it is crucial to implement a suitable pipeline Non-Destructive Testing (NDT) system so that it can avoid wastage of resources, unscheduled shutdowns and prevent catastrophic pipeline failures.

Latest advancements in sensing technology have yielded a system that is able to continuously monitor pipeline segments using LRUT [5]. This system is able to monitor pipelines over a distance of several hundreds of meters from a single location using a ring of ultrasonic transducers permanently fixed onto the pipe. LRUT was specifically designed for inspection of Corrosion Under Insulation (CUI) and has many advantages over other NDT techniques which have seen its widespread use in many other applications [6]. It is also able to detect both internal and external corrosion which makes it a more efficient and cost-saving alternative. With the recent developments in permanent mounting system using a

* Corresponding author. Tel.: +605 2490500.
E-mail addresses: keyx1nak@nottingham.edu.my (N.A. Akram), Dino.Isa@nottingham.edu.my (D. Isa), Rajprasad.Rajkumar@nottingham.edu.my (R. Rajkumar), lamhong.lee@qiup.edu.my (L.H. Lee).

http://dx.doi.org/10.1016/j.ultras.2014.03.017
0041-624X/

Table 1
PHMSA all reported pipeline incidents by cause 1993–2013 (December 6, 2013). (Source: http://primis.phmsa.dot.gov).

Reported cause of incident	Number	%
Corrosion	1926	18.40
Excavation damage	1950	18.60
Incorrect operation	742	7.10
Material, welding, equipment damage	2818	26.90
Natural force damage	714	6.80
Other outside force damage	765	7.30
All other causes	1535	14.60
Totals	10,450	100.00

special compound, the ability to perform a continuous monitoring system has now become a reality [7]. Although this technology exists, there is currently no continuous monitoring and prediction software available that can automatically make real-time decisions on the status of the pipeline [8].

The challenge for continuous monitoring using SVM is that the conventional SVM only works with batch data. It processes all previous data to create a decision model [9]. It will need more memory and computing power as new data is acquired [10]. Hence this technique is not suitable for a continuous monitoring system which runs non-stop and relies on real-time data. Active incremental SVM is more suitable as it will only process newly acquired data. It discards all previous data and as a result, the memory demand is consistent at all the time. However, it may sacrifice some valuable information in the discarded data and may affect the classification accuracy [11–13]. Therefore, the proposed approach is designed to work in active incremental mode in order to retain high classification accuracy. Fig. 1 illustrates the difference in data points usage in the training phase of conventional SVM compared to active incremental SVM, in different stages of data acquisition.

This work proposes a real-time approach that uses SVM classifier equipped with the active incremental training approach. More detailed explanations on the proposed algorithm are presented in the following sections as well as details of the experimental setup and data acquisition. The experimental results show that the proposed technique achieves low time consumption and promising classification accuracy compared to when the SVM classifier is trained in batches.

2. Support Vector Machine

SVM has been extensively used in pattern recognition and regression. It has comparable computational efficiency and excellent generalization capabilities. SVM was originated from Vladimir Vapniks̈ in 1970, with the idea of the Structural Risk Minimization (SRM) [14]. SVM performs recognition by forming a hyperplane that separates two classes with a maximum distance as shown in Fig. 2. These points are called support vectors. In many cases, data are not linearly separable in the input space, thus non-linear transformation is applied [15].

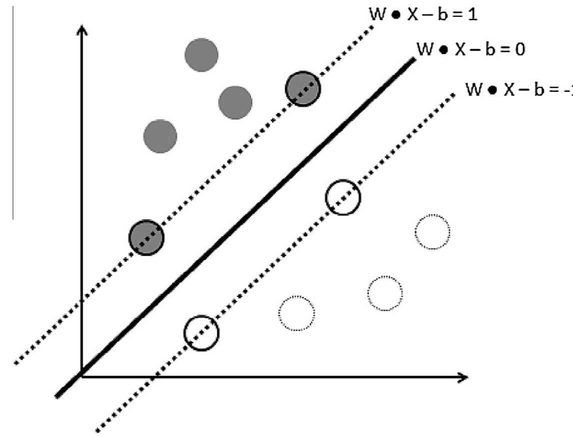

Fig. 2. Two classes of dataset.

2.1. SVM as pattern classifier

SVM most often performs similar to a human's brain in pattern classification. It is able to identify and distinguish an object by its shape, feature, appearance, etc. [16–18]. For example, given a set of patterns as illustrated in Fig. 3, there are a few ways to draw a separating line for these two classes as shown in Fig. 4.

The SVM will find a hyperplane having a maximum margin to separate the training patterns in both classes [19]. Closest sample points to the hyperplane are defined as 'support vectors' (SVs). SVM will store only the SVs in the classification phase (see Fig. 5).

2.2. SVM formulation

Based on Fig. 2, it illustrates that the hyperplanes of the SVM are represented by Eqs. (1)–(3). The optimal separating hyperplane of the SVM is defined as Eq. (1). The training vectors are linearly separated by a hyperplane. The hyperplane is represented by a linear equation in Eq. (1) where **w** is normal to the hyperplane and b is the bias.

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \quad (1)$$

The training vectors belong to different classes which are +1 class and -1 class. Eqs. (2) and (3) represents the supporting hyperplane.

$$\mathbf{w} \cdot \mathbf{x} + b = +1 \text{ for class } +1 \quad (2)$$

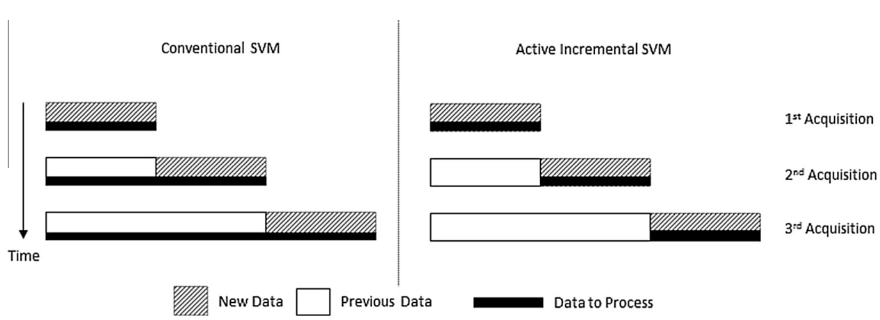

Fig. 1. Comparison between training process of conventional SVM and active incremental SVM.