ELSEVIER

Journal of Materials Processing Technology 124 (2002) 305–310

Journal of Materials Processing Technology

www.elsevier.com/locate/jmatprotec

AI-based condition monitoring of the drilling process

B. Brophy[a,*], K. Kelly[b], G. Byrne[a]

[a]Mechanical Engineering Department, University College Dublin, Belfield, Dublin 4, Ireland
[b]Department of Mechanical and Manufacturing Engineering, Trinity College, Dublin 2, Ireland

Accepted 10 April 2002

Abstract

With increasing competitive pressures, manufacturing systems in the automotive industry are being driven more and more aggressively. The pressures imposed on the processes and lack of system 'slack' have led to increased use of tool condition monitoring (TCM) systems. In parallel, there has been wide-ranging research in academia. However, a closer examination shows that there has been very little migration of this research into industrial practice. Furthermore, the success of industrially deployed monitoring systems has been poor. It has been suggested that a significant factor behind both these phenomenon has been the 'difficult' environment in which such systems must operate; an environment where they are subject to many stochastic influences, ranging from ambient conditions, to user input, to workpiece consistency.

Neural networks (NNs) have found increasing favour in manufacturing systems research because of their ability to perform robustly in noisy environments. Almost all the applications of this technology in TCM have been in the detection/prediction of tool wear. From an academic standpoint, it may be speculated that the lack of focus on breakage and missing tool detection has been due to the relatively trivial nature of detecting such anomalies in the laboratory environment. However, detection in the production environment is compromised by a wide range of factors, which can give rise to false alarms when such strategies are transported from laboratory conditions. In this paper, data from a real manufacturing process is used to demonstrate the potential application of NNs to the task of anomaly detection in the production environment.

Keywords: Monitoring; Drilling; Intelligent

1. Introduction

1.1. Tool condition monitoring (TCM)

The demand in the manufacturing industry to reduce production costs has driven major manufacturers to automate many operations. Productivity increases are often limited by the requirement that such automated operations have a high level of reliability. Part scrappage or damage to machines as a result of broken tooling, tool wear, tool collisions, inadequate swarf removal, or coolant malfunction, among others, must be avoided. Several strategies are available to manufacturers to reduce the possibility of such failures. The two main options, apart from returning to manual observation of each operation, are to run the machine tool well below its maximum capacity, or to introduce some form of feedback (representing the quality of the process to the system). A more specific case of this monitoring of the process is where attention is focussed on the state of the cutting tool. These are termed 'tool condition monitoring' systems.

* Corresponding author.

Broadly speaking, there are two possible methods of determining tool condition. In the direct method, tool condition is measured directly in situ. Measuring devices based on inductance, capacitance, vision, radiation, or pneumatics can be used to measure the level of wear [1]. While direct methods tend to be accurate, they are more complex and often not applicable to real machining environments. They are also usually only implementable between operations meaning that breakage or excessive wear can only be detected post-process [1]. Indirect methods involve the measurement of some phenomenon related to tool wear or breakage. Commonly, the measurement of cutting force, torque, temperature, vibration, spindle motor power, feed motor power, and strain are used to indirectly indicate the level of tool wear. As the measured phenomenon often varies with process conditions it can be difficult to find a good correlation between tool wear and the sensor signal over the full range of operating conditions. Indirect methods do, however, offer the advantages that they: can continuously monitor the process, are less complex than direct methods, and are applicable in industrial environments. For these reasons, the majority of the research carried out in TCM has concentrated on indirect methods.

0924-0136/02/$ – see front matter © 2002 Published by Elsevier Science B.V.
PII: S0924-0136(02)00201-7

The correct choice of sensor(s) is essential to the operation of any monitoring system [2]. If there is poor correlation between the sensor signal and the tool condition it is unlikely that correct classification of tool state can take place. Deterioration of drill condition manifests itself in a change in the cutting forces and so monitoring these forces should provide a good indication of tool state. In drilling, the measurement of torque, thrust force and transverse strain gives information on the cutting forces. Torque can be measured directly with a torque sensor or indirectly through measurement of the spindle motor power and the spindle speed. It is worth noting, however, that spindle motor power can be used as a valid indicator of tool condition only when the cutting process consumes a significant proportion of the total spindle power [1]. Non-symmetric wear of a drill or drill asymmetry can be identified through the analysis of the vibration signals [3].

In summary, the majority of researchers investigating TCM in drilling have measured torque, thrust force, vibration and strain [3–6]. Other researchers have measured temperature and sound [7] and acoustic emission (AE) [8–10].

Applications of TCM in industry have depended mostly on robust and reliable sensor signals such as force, power and AE. The monitoring strategies used by TCM system manufacturers are largely based on limits and enveloping functions. Most require the capture of a reference pattern (the 'teach-in' signal), where the sensor signal is recorded and stored as being representative of a typical machining operation. Certain limits are applied to the reference signal based on the heuristic knowledge of the system manufacturer taking into account the individual set up. During subsequent machining operations the recorded signals are compared with the taught-in signal and appropriate action is taken by the TCM system based on the result. The action of "teach-in" may be required for every new tool, new part and new material. Hence this methodology can be very tedious and time consuming, in particular when TCM is applied to flexible machining systems.

1.2. Intelligent TCM systems

Research into the area of TCM has proceeded apace, although advances in research have typically not found their way to practically applied monitoring systems. There are doubtless many factors behind this, although principal among them would appear to be concerns about reliability. This hesitance is understandable in the light of the controlled conditions under which most strategies have been tested [11]; typically off-line and with a very limited range of operating conditions. Further to this point, there are many stochastic influences on the process in production environments (e.g. variations in workpiece composition) which are difficult, if not impossible, to replicate during laboratory research.

The success of any TCM system is dependent on two factors, the quality of the data acquired by the sensors and the diagnosis algorithm used to analyse the sensory information and determine tool state. Intelligent models for tool breakage detection have gained more and more attention in recent years largely because they can better approximate the correct mapping relationship between inputs and outputs of a dynamic system directly (including an implicit accounting for the stochastic influences mentioned above), whereas physical models require the derivation of very complex mathematical equations involving quantities that are difficult to ascertain [12]. Dimla et al. [11] have referred to the efforts of many researchers in recent years to move from the 'predictive' type of system model to an intelligent discriminator type of system. Examples of such 'intelligent discriminators' are expert systems, fuzzy sets, and neural networks (NNs). NN-based algorithms are perhaps the most suitable for integrating multiple sensor information due to their strong ability to describe the highly non-linear characteristics of machining processes, superior learning, noise suppression, and parallel computation abilities.

Two main classes of network architecture may be identified, namely the 'supervised' and 'unsupervised' types. This classification refers to the method used to train the network, i.e. the means by which the reference/inference information is encoded within the network structure. Supervised learning requires the intervention of the trainer to provide feedback to the network to further the training progress. Unsupervised learning proceeds by assembling or grouping the training data into 'clusters', i.e. the network produces a topologically ordered map of the input data. In both cases, the network is typically used (after the training phase has been completed) as a classifier.

Many different network architectures have been applied to the TCM problem, spanning both the supervised and unsupervised paradigms [11]. The most popular network structure has been that of the multi-layer perceptron trained by back-propagation. Other methods, including Kohonen self-organising maps and adaptive resonance theory have also been applied. The vast majority of work presented in the literature has focussed on the identification of tool wear, and reported success rates of 95% are fairly typical. Tansel et al. [13–15] have looked at recognition/prediction of tool breakage, particularly in micro-drilling.

What the research has typically had in common, however, was feature extraction from a large volume of sensed data (often from several sensors, i.e. so-called 'sensor-fusion'). Dimla et al. [11] have identified four typical phases in this feature extraction:

- Sensor selection,
- Primary transformation, via FFT or wavelet analysis,
- Secondary transformation, through derivative or integral operations,
- Tertiary transformation, through statistical analysis such as mean, standard deviation, kurtosis, etc.

Of course, not all researchers have used all four stages, but it is almost invariably the case that statistical parameters extracted from the (processed) sensor data have been used in the training of the network.

The work reported here (and in an earlier publication [16]) differs substantially in respect of the fact that a single raw sensed signal (in this case the spindle motor power) is passed directly to the network.

2. Network development

The work described in this paper was initiated as part of the COMPRO project (see Acknowledgements). This was a collaborative research effort in the area of TCM, involving both industrial concerns and research institutes. It was funded by the European Commission under the BRITE-Euram funding framework. During the project, extensive testing and measurement was carried out by the authors at the sites of several of the industrial users. The use of NNs for anomaly detection made up one important part of the COMPRO project.

Section 2.1 describes the development of the NN during the COMPRO project, and Section 2.2 summarises some of the results obtained. The network was trained with data acquired from an automated drilling operation over a 3-week period. The network was then tested on data from different machining operations but gathered during the same period. However, Section 2.3 describes a new test regime for this network. Subsequent to the COMPRO project, data was gathered by a different system over a period of several months. The difficulties that this new data posed to the NN are explained in Section 2.3 and there is a summary of the classification success for these later tests in Section 2.4.

2.1. Initial set up and testing of network

The aim of the work reported here was to develop a NN capable of detecting anomalies (such as tool breakages and missing tools) in industrial drilling processes. Operations were classified as being either 'normal' or 'abnormal' on the basis of the spindle power signal from the machine tool. The goal therefore was to replicate the judgement of an experienced operator using the network, with potential benefits for the monitoring of unmanned machining operations. The spindle power signal was acquired for a full drilling operation and passed to the network. The network incorporates two distinct stages (Fig. 1):

Feature extractor. The first stage of the network uses unsupervised learning to extract the so-called 'principal components' from the raw power signal. It does this without prior definition of what these principal components are. The reduction in the number of data points making up the input, to the number of principal components extracted, is roughly 10:1.

Classifier. The principal components extracted in the first stage form the input to the second stage. This stage learns through supervised learning (back-propagation) to classify the principal components from the first stage as being representative of either a normal or abnormal operation.

The two-stage network was first presented with a set of 'training' files along with a set of known output classifications. When the network was adequately trained, a set of 'testing' files was introduced to see how well the network classified files it had not previously encountered.

The machining operation under examination was gun-drilling at a station in a transfer line that was machining diesel truck engine blocks. The bore of the drilling operation was 12 mm and the duration was 185 s. The workpiece material was cast iron.

Ideally, the network would learn to classify the power signals as representing either 'normal' or 'abnormal' drilling operations. (An abnormal signal would result from a missing tool or a tool breakage during machining.) The network should thus be trained on a representative set of signals from both normal and abnormal operations. However, there was only one abnormal signal (due to a tool breakage) within the entire data set for the two processes. In order to test the network, a number of abnormal signals had to be simulated or generated.

'Missing Tool' signals were easily simulated as flat line power traces. For 'breakage' signals, the shapes were chosen to imitate breakage signals encountered by experienced operators, on this and similar processes. Both 'high-speed steel' type (exhibiting a spike in the power signal when breaking) and 'tungsten carbide' type (typically no spike on breaking) breakages were simulated. Care was also taken to ensure that there was a variation between network training and testing sets in the size, timing and type of breakage signals used. A sample of these breakage signals is shown in Fig. 2.

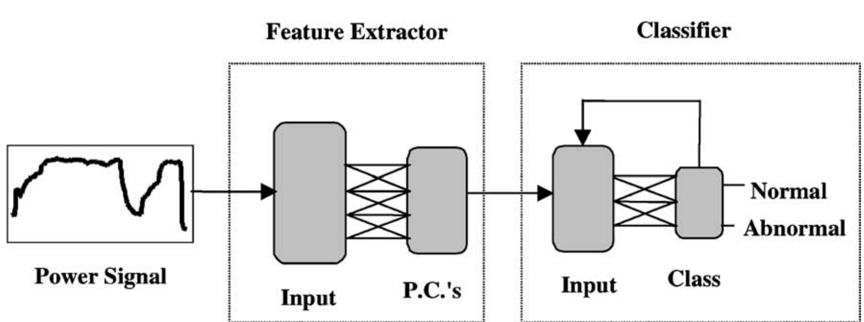

Fig. 1. Schematic of NN.

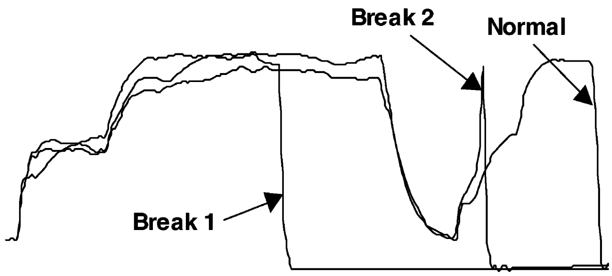

Fig. 2. Sample of breakages.

Table 1
Classification success of network for first series of tests

Training	
No. of normal	79
No. of abnormal	6
Training time (min)	15
Testing	
No. of normal	79
Correctly classified	79
No. of abnormal	6
Correctly classified	6

2.2. Results of initial tests

The classification success of the network is summarised in Table 1. All operations were classified correctly.

The network was able to recognise the basic pattern of the power plot from the drilling operation despite considerable variation in the mean level of the signals. The mean level of the largest signal was nearly two-and-a-half times the value of the smallest one. There was also a large amount of in-process noise as well as many aberrations in the shape of the signals. Such variations would pose considerable problems to a conventional level-based monitoring system. It can also be seen that the training time (around 15 min) is relatively short.

2.3. Long term unsupervised testing

Although the network performed well on this real industrial data, the training and test data sets were gathered during the same time period. It could therefore be contested that the network was being trained with the test data in mind. In conjunction with this, the breakages had to be simulated as there was only one actual breakage signal recorded during the 3-week acquisition period. What was now required was a longer-term test run. A recording module was left attached to the process over a period of several months. It was hoped that there would be more process variability over this longer time period and that this would provide a sterner test for the network. It was also hoped that there would be some breakages during this time.

Data was collected for 3 months. In that time, the power signals from over 2900 drilling operations were recorded. There were some distinct differences between this data set and that used to test the network initially. Data was acquired using a different acquisition module with a much poorer resolution than that used before. This introduced noise in the form of quantisation error into the new test data. Fig. 3 shows two power signals from this new data set. The signals both came from normal machining operations. The quantisation noise can be clearly seen on both plots. The upper plot also displays a 'hump' near to the start of machining. This increase in power may have been caused by a build up of swarf that later cleared. There were many such aberrations in the shapes of the power curves in this new data set that were not present in the original data. This posed new, unforeseeable challenges to the network that had been trained on the original data set.

Fig. 3. Two 'normal' power signals from the new data set.

There were no breakages recorded during the 3-month trial. However, there were several operations that displayed abnormal spindle power plots. Two such plots are shown in Fig. 4. The solid line plot came from an operation where machining was aborted (probably due to a breakage of one of the other drill bits at this machining station). The broken line plot came from an operation which immediately followed an aborted operation and for which the first part of machining had already been carried out.

These abnormal signals were used as the anomalies for this 3-month test batch. The technician in charge of the transfer line had mistakenly thought that many of these operations were actual breakages. The exact reason for each of these abnormal curves was only discovered at a later stage. Thus, although it was once again impossible to test the network with real breakages, at least the signals used were considered genuinely anomalous by an experienced technician and worthy of further examination. The network could

Fig. 4. Two 'abnormal' power signals from the new data set.

Table 2
Classification success of on data gathered over 3-month period

Training	
No. of normal	79
No. of abnormal	6
Training time (min)	15
Testing	
No. of normal	27
Correctly classified	27
No. of abnormal	6
Correctly classified	5

be considered successful if it were able to distinguish these signals from the normal signals described earlier.

2.4. Results of long term tests

The trained network from the first series of trials was used. This network was not retrained with any of the new data. Of the new data set, power plots from 27 normal operations and six abnormal operations, selected representatively from the new data, were used for testing. The classification success with this set of data is summarised in Table 2.

The output of the network is a probability of whether the operation represents either of the two states; normal or abnormal. The single incorrect classification made was on the operation (dashed line) shown earlier in Fig. 4. This operation was the completion of one that had been begun before. The reason that it was not detected as an anomaly is that there was no operation that resembled this one in the training set. It is not unreasonable, therefore, that the network would guess incorrectly that this was supposed to be an abnormal signal.

3. Discussion and conclusions

A two-stage NN has been used to detect anomalies in the drilling process. The network has been used to classify drilling operations as 'normal' or 'abnormal' (tool breakage or missing tool). This network is distinguished by the fact that it uses the spindle power signal (acquired over all or part of the operation) as its input, rather than statistical features extracted from this signal.

A further distinguishing feature of the work reported here is that the data used in the training and testing of the network were acquired from a working industrial process. High classification success has been achieved with short training times. The network was originally trained and tested on distinct sets of data obtained contemporaneously, as reported previously [16]. This paper has also examined the performance of the network (without re-training) on data acquired using a different logging module and over a longer (and separate) period of time. Furthermore, the recurrent difficulty found with currently applied strategies (where the limits required to ensure recognition of breakage also results in a large number of false alarms) has been overcome, with no false alarms being encountered in either of the test sets.

Although there was only one actual breakage signal produced by either of the test runs, it is worth noting that the abnormal signals used in the second test run were originally identified by an experienced operator as resulting from breakages. Given that the goal of an intelligent system must to be replicate sound human judgement, the ability of the network to also recognise these signals as abnormal shows that it does act as an intelligent discriminator. In the case of the single operation that it failed to classify as abnormal, it should be pointed out that no such signal (where machining resumed on a previously aborted operation) had been encountered in the training set. Periodic re-training of the network to include such anomalous signals could redress this difficulty; in much the same way as a human operator learns from experience. Alternatively, incorporation of information from the preceding machining operation (e.g. total energy consumed) could solve the problem.

A limitation of the approach detailed in this paper is that it requires the full signal before a classification is made. For many operations this is acceptable, but for operations of very long duration, or where undetected breakage is catastrophic, additional precautions and techniques are required. Some such methods, involving dynamically evaluated networks using partial data are explored in detail elsewhere [16]. A further promising avenue for research is a hybrid approach combining elements of NN-based strategies with more conventional limit-based algorithms. This is currently under investigation.

Acknowledgements

Funding for much of this work was provided by EU BRITE project No. BE96-3565, COMPRO (Compensation for Stochastic Deviations in Process Monitoring for Flexible Production Systems) and is gratefully acknowledged. The authors are also grateful for the co-operation of their partners in this project (Robert Bosch GmbH, Hüller Hille GmbH, Ideko, Scania AB, Comau SpA, Kistler Instrumente AG, Artis GmbH and FhG-IPK Berlin).

References

[1] G. Byrne, D. Dornfeld, I. Inasaki, G. Ketteler, W. König, R. Teti, Tool condition monitoring (TCM)—the status of research and industrial application, Ann. CIRP 44 (2) (1995) 541–567.

[2] L. Kang-Ning, L. Cheng-Jen, An intelligent sensor fusion system for tool monitoring on a machining centre, Int. J. Adv. Manuf. Technol. 13 (1997) 556–565.

[3] S.Y. Hong, J. Ni, S.M. Wu, Analysis of drill failure modes by multi-sensors on a robotic end effector, Japan/USA Symposium on Flexible Automation, Vol. 2, ASME, 1992, pp. 947–956.

[4] A. Noori-Khajavi, R. Komanduri, Frequency and time domain analysis of sensor signals in drilling I & II, Int. J. Mach. Tools Manuf. 35 (6) (1995) 775–815.

[5] T.I. El-Wardany, D. Gao, M.A. Elbestawi, Tool condition monitoring in drilling using vibration signature analysis, Int. J. Mach. Tools Manuf. 6 (6) (1995) 687–711.
[6] T.I. Liu, W.Y. Chen, E.J. Ko, Intelligent recognition of drill wear states, J. Mater. Eng. Perform. 3 (4) (1994) 490–495.
[7] P. Haili, C. Yubao, O. Elsayed, Monitoring methods of tool wear in a drilling process, in: Proceedings of the Second Industrial Engineering Research Conference, Los Angeles, 1993, pp. 380–384.
[8] W. König, G. Ketteler, T. Klumpen, Monitoring turning, milling, drilling and grinding processes using acoustic emission signals, Progress in Acoustic Emission VII, The Japanese Society for NDI, 1994, pp. 619–624.
[9] D.A. Dornfeld, M.F. DeVries, Drill wear and Fracture Monitoring Using Acoustic Emission, Presented at CIRP STC ‘‘C’’ meeting, Paris, 1992.
[10] Y. Saijo, A. Noh, A method of detection of drill breakage in drilling, in: Proceedings of the Autumn Annual Meeting of JSPE, 1978, pp. 11–12.
[11] D.E. Dimla Jr., P.M. Lister, N.J. Leighton, Neural network solutions to the tool condition monitoring problem in metal cutting—a critical review of methods, Int. J. Mach. Tools Manuf. 37 (9) (1997) 1219–1241.
[12] M.S. Lan, D. Dornfeld, In process tool fracture detection, Trans. ASME, J. Eng. Ind. 106 (2) (1984).
[13] I.N. Tansel, C. McLaughlin, On-line monitoring of tool breakage with unsupervised neural network, Trans. NAMRI/SME 19 (1991) 364–370.
[14] I.N. Tansel, O. Rodriguez, C. Mekdeci, Detection of tool breakage in micro-drilling operations with RCE neural networks, ASME PED 47 (1) (1992) 91.
[15] I.N. Tansel, Monitoring micro-drill operations using neural networks, Key Eng. Mater. 138–140 (1998) 575–592.
[16] B. Brophy, K. Kelly, G. Byrne, Anomaly detection in drilling using neural networks, in: Proceedings of the Second International Conference on International Manufacturing System, Leuven, Belgium, September 1999, pp. 779–786.