Milling-Tool Wear-Condition Prediction with Statistic Analysis and Echo-State Networks

Guodong Wang & Radu Grosu
Vienna University of Technology, Vienna, Austria

ABSTRACT: Tool wear is the most commonly observed and unavoidable issue in metal milling. The worn or damaged cutting tools will cause materials loss and machines shut-down. To tackle this problem, we propose a new method for predicting the wear condition of end-milling tool. First, we adopt statistic-analysis techniques to analyze the collected data. Second, we select interesting features based on Pearson-correlation coefficient (PCC). Finally, those features are applied as inputs to the so called echo-state network to predict subsequent tool wear condition. The experimental results and theoretical analysis both demonstrate that the proposed method performs better than naive feed-forward neural networks (FFNN) and time-series neural networks (TSNN).

1 INTRODUCTION

Milling is the machining process of using rotary cutters to remove material from a workpiece progressing (or feeding) in a direction and an angle w.r.t axis of the tool (García-Nieto 2015). It covers a wide variety of different operations and machines varying from small individual parts to large, heavy-duty gang milling operations. Due to the up-streaming property of the drilling process, a smooth tool surface always implies a high quality of production,otherwise, the products are unacceptable. For instance, a dull tool may tear the surface and decrease the fatigue resistance of the workpiece. A worn tool can cause more friction which will increase the cutting temperatures and lead to resource waste. Tool wear monitoring and prediction has attracted a considerable amount of research attention in the past decades, as tool wear has a great effect on the final products as previously mentioned. Therefore, efficient solutions should be proposed to solve this problem. There are many types of tool wear, for example, rounding wear of the cutting edge, crater wear on the rake face, friction caused flank tool wear and so forth. Several parameters for causing wear have been investigated, e.g., the cutting speed, feeding rates, cutting depths.

In recent years, plenty of methods have been proposed to help to produce high quality products. Generally, these methods are mainly using statistic techniques, signal-processing methods, neural networks and so on. The representative methods include: Ai proposed to use acoustic spectrum (Ai et al. 2012) whereas De Jesu used current signal to analyze the tool wear condition (De Jesu et al. 2003). (Ghosh etal.2007) and (Palanisamy et al.2008) proposed to use Feed-Forward Neural Networks (FFNN) for monitoring tool wear. On the other hand, (Mosallam et al.2014) introduced a method for trends extraction from multidimensional sensory data that considers time. This method is based on extracting successive features from machinery sensory signals. According to the literature, the aforementioned methods work fine for the tool-wear issue. However, the the problem is that part of the methods only focused on single feature and ignored other ones like spindle or table vibration which lead to unreliable monitoring results while others did not consider the time variation issue. In fact, the machining process is a time series procedure as the post tool condition happens based on its prior states. In contrast, in this paper we regard tool wear changing process as a time series procedure generating sequential data points in one time interval. At each time step t of the interval, one data point will be generated which may depend on its previous $t-1$, $t-2$, or $t-n$ data points.

To achieve a sustainable manufacturing, we propose in this paper a novel method for predicting milling-tool healthy states based on echo-state networks (ESN). These are a new type of recurrent neural networks that have the ability to memorize. The major advantage of ESN is the simplicity of their implementation and training. We used statistical analysis techniques for extracting mutil-sensor features of machining tool. Since cutting tools are gradually changing over time, ESN is adopted as the predictor of next healthy state.

The main contributions of this paper can be briefly summarized as follows:

(1) A tool wear prediction method is proposed based on statistic analysis techniques and echo state network.

(2) A numerical comparison with existing methods (TSNN and FFNN) is provided, showing the prediction accuracy of our technique.

The rest of the paper is organized as follows. In section 2, we introduce the experimental data set and the signals-collection setup. In section 3, we propose our main algorithm for tool-wear-states prediction. In section 4, we introduce the configuration of ESN and present the simulation results. We end the paper with concluding remarks and research directions.

2 EXPERIMENT DATA SET AND SYSTEM STRUCTURE

One specific industrial milling machine is used for the experiments. The machine operates to get an experimental data-set under various conditions. To collect the machining-processes-generated data, three different types of sensors are installed around the tool at several positions. Those sensors are acoustic emission, vibration and current sensors.

The generated data is organized in a 1×167 matlab struct array with fields as shown in (Goebel 1996). The experiment is divided into 16 different study cases. Each case with various number of runs. Tool flank wear is measured at irregular time intervals. For each measurement, one has to stop the whole machining process.

We acquired the experiment result data set from NASA data repository (Agogino & Goebel 2007). (Goebel 1996) gave the concrete introduction of the experiment setup. It encompasses the characteristics of the used physical devices, for instance, the size of workpieces, types of materials and inserts. It also describes the various experimental conditions e.g. position of installed sensors, feeds rates, cutting speed and depths. The collected signals from all sensors are preprocessed before entering the industrial computer. The preprocess consists on amplifying and filtering those data, then feeding processed data into two RMS devices. The current signal of spindle motor is directly fed into computer without further RMS process.

The major reason for using device RMS is that it can smooth the signal and make it more accessible for signal processing. Moreover, RMS is proportional to the energy contents of the signal (Goebel 1996), according to the Formula 1. The whole process is depicted in Figure 1. In Formula 1, $\Delta T = 8.00$ ms, sampling rate is 250HZ, and $f(x)$ is the signal function.

$$RMS = \sqrt{\frac{1}{\Delta T}\int_0^{\Delta T} f^2(x)dt} \tag{1}$$

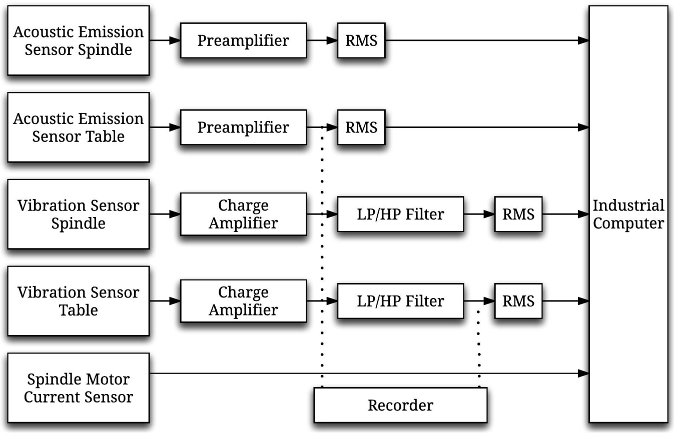

Figure 1. Experimental setup.

3 THE PROPOSED METHOD

In this section, we first introduce a way to select appropriate features for the data set. Then give a brief introduce of the ESN model and its key global parameters (spectral radius, sparsity of reservoir, size of ESN, input scaling and leaking rate). Finally, the specific steps for implementing the proposed method is listed out.

3.1 *Features extraction and selection*

The collected raw data always contains noise, abnormal data points and other unexpected problems. To obtain a good data set for the analysis, the most valuable and decision-relevant features should be extracted and selected. Based on studies (Dong etal. 2006) (Wang et al. 2013.), we consider the following features extracted from raw data which are listed in Table 1. CF is the abbreviation of crest factor, P/AR is the peak to average ratio and skewness. Different features represent various information about tool heath states, for instance, RMS is the measure of a varying quantity which is also related to the energy of signals. Kurtosis indicates the plus of signals and Skewness characterizes each degree of asymmetry of the distribution around its mean. Naturally, not all features are valuable, but it is hard to decide which one is more sensitive to the tool wear. A good feature should present consistent trends with defect propagation (Wang et al. 2013.).

In this paper, we use PCC, which measures the independence of two or more random variables, to rank features. The way of computing PCC is shown in Formula 2.

Table 1. Summary of extracted features

Features	Expression
Mean value	$\overline{x} = \frac{1}{n}\sum_{i=1}^{N} \lvert x_i \rvert$
RMS	$X_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i)^2}$
Variance	$X_{var} = \frac{1}{N}\sum_{N}(x_i - \overline{x})^2$
Maximum	$X_M = max(x_i)$
CF	$C_F = \frac{X_M}{X_{RMS}}$
Kurtosis	$X_{KURT} = \frac{1}{N}\sum_{N}(\frac{1}{\sigma}(x_i - \mu))^4$
P/AR	$X_{PAR} = \frac{X_M}{\overline{X}}$
Skewness	$X_{SKEW} = \frac{1}{N}\sum_{N}(\frac{1}{\sigma}(x_i - \mu))^3$

$$PCC = \frac{\sum_i (x_i - \overline{x})(z_i - \overline{z})}{\sqrt{\sum_i (x_i - \overline{x})^2 (z_i - \overline{z})^2}} \quad (2)$$

3.2 *ESN model*

ESN is a new typology of recurrent neural network (RNN), first proposed by (Jaeger & Haas 2004). It has a naive way of generating the neural network and uses a linear regression algorithm to train the network. Recently, ESN has attracted a big amount of research efforts. It has plenty of successful applications, especially in dynamic pattern recognition, robot control and time series prediction (Jaeger & Haas 2004) (Shi & Han 2007). The typical ESN follows the structure depicted in Figure 2. It has three different layers which are the input layer, the reservoir and the output layer. For a given machining system, the input is:

$$u(n) = [T(n), C_a(n), C_d(n), V_t(n), V_s(n), A_t(n), A_s(n)]' \quad (3)$$

where T is the running time of milling system, C_a, C_d, V_t, V_s, A_t and A_s are AC/DC motor current, table/spindle vibration and acoustic emission from table/spindle, respectively. A target output $y^{target}(n) = VB \in R^{N_y}$ is given, which is known as flank wear, and the actual output will be $vb(n)$, N_{vb} is the output dimensions, here $N_{vb} = 1$. W^{in} is the weight matrix from input to reservoir and W is the randomly generated connection matrix from reservoir to reservoir. $W^{out} \in R^{out}$ is the connection weight matrix from reservoir to output.

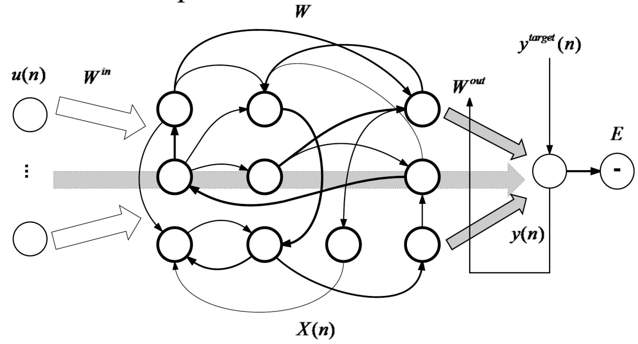

Figure 2. Basic ESN structure.

In the initial phase, ESN randomly generates a hidden-layer connection which is called a reservoir. The input-to-reservoir connection-matrix and reservoir inner-connection will be randomly generated and kept fixed during the training phase.

Generally, ESNs are applied to supervised temporal machine learning tasks. To train the ESN, $y^{(n)}$ should close enough to y^{target} according to desired accuracy. Hence, we can minimize the error measurement $E(vb, vb^{target})$, typically by using the Root-Mean-Squired-Error (RMSE) to achieve above goal, the formula is:

$$E(vb, vb^{target}) = \frac{1}{N_{vb}}\sum_{i=1}^{N_{vb}}\sqrt{\frac{1}{T}\sum_{n=1}^{T}(vb_i^n - vb_i^{target})^2} \quad (4)$$

Once the input is fed into ESN, the reservoir state $x(n)$ will be updated with following formulas:

$$\widetilde{x} = f(W^{in}[T(n), C_a(n), C_d(n), V_t(n), V_s(n), A_t(n), A_s(n)]' + Wx(n-1) + W^{out}x(n)) \quad (5)$$

$$x(n) = (1-\alpha)x(n-1) + a\widetilde{x}(n) \quad (6)$$

Where f is the activation function and α is the leaking rate which is used as the memory loss factor of ESN for the past states. Here, we choose a sigmoid model as the activation function. The output layer is defined as a linear combination of the reservoir states and inputs:

$$vb(n) = g(W^{out}[T(n); C_a(n); C_d(n); V_t(n); V_s(n); V_t(n); V_s(n); A_t(n); A_s(n); x(n)]) \quad (7)$$

Where [:;:] stands for column vector (or matrix) concatenation, g is the output activation function. The computing performance of an ESN depends on the following key parameters: The spectral radius, which is used for generating the echo state and one needs to make sure that the maximum eigenvalue of spectral radius is always less than 1. The sparsity of reservoir, which decides the mount of connection among neurons. The size of ESN, which is the amount of neurons embedded in the reservoir.

Generally, with a bigger number of neurons, an ESN performs better while it also increase the complexity of ESN at the same time. Input-scaling maps the input into the range of neuron activation function. Leaking-rate can be regarded as the speed of the update dynamics of the reservoir discretized in time (Lukosevicius 2012).

In summary, compared with classical RNNs, the advantages of ESNs are:

(1) Randomly generated reservoir connections support rich-dynamics internal states for the prediction of the output.

(2) Randomly generated connection matrix from input to reservoir and from reservoir to reservoir, which are kept fixed during training process, overcomes the low computing efficiency issue of RNN.

(3) For complex non-linear problems, the linear combination of reservoir states can always get accurate solutions.

3.3 *Specific steps for implementation*

In this section, we introduce our method for tool wear prediction. The main idea is to use statistical techniques for extracting useful information and deploying ESN to estimate tool healthy states. The proposed method is summarized as follows:

Step 1: Collect raw data set from sensors U_n and make appropriate configuration for ESN, W^{in} and for W.

Step 2: Clean the data set: remove and make up the missed or abnormal value.

Step 3: Based on the cleaned data from second step, use statistic analysis techniques to extract the most valuable and relevant information for decision: we adopt 8 different features.

Step 4: Use Pearson correlation coefficient (PCC) to rank those features for each type of signal.

Step 5: Use a part of the processed data as inputs to train ESN: one can manually get the parameters for ESN.

Step 6: Apply the rest of the data as inputs to ESN and predict the coming states.

Step 7: Return tool wear condition VB.

4 EXPERIMENTS

4.1 *Configuration*

According to the description of the experimental-data set, we consider one running case as one time series process. For each case, one specific ESN is designed to predict the tool wear condition. The configuration of each ESN is manually acquired. Six different running cases are used to train ESN and validate the accuracy of the prediction. The selected data set are labeled as case 1, case 2, case 3, case 11, case 12 and case 13, respectively.

$$MSE = \frac{1}{N_{vb}} \sum_{i=1}^{N_{vb}} (vb_i(n) - vb_i^{target}(n))^2 \tag{8}$$

To evaluate the performance of the proposed algorithm, we also compared the simulation results obtained through ESN, FFNN and TSNN methods. In the experiments, the flank wear VB is measured as a generally accepted parameter for evaluating the tool wear condition. VB is observed during the experiments and measured as the distance from the cutting edge to the end of the abrasive wear on the flank face of the tool. The measurement is Formula 8.

4.2 *Simulation results*

To reduce the dimension of inputs, PCC is applied for ranking features, the one with the top score is selected for future prediction. The analysis results for different study cases are listed in Table 2. Because of the limitation of pages, we only list out part of PCC ranking results of 6 study cases. We implement the proposed algorithm in MATLAB, for all simulations, each running case uses same training and validation data set.

Table 2. PCC ranking.

smcAC	smcDC	Vib_table	Vib_spindle	AE_table	AE_spindle
CASE 3					
0.815	0.9467	-0.8006	-0.4802	0.9547	0.8333
0.9514	0.9521	-0.826	-0.4687	0.9541	0.826
0.9574	0.9145	-0.8089	0.4894	0.898	0.6431
0.9232	0.9466	-0.86	0.7322	0.9228	0.6618
0.03612	0.3632	-0.7639	0.7641	0.1259	-0.1147
-0.08484	-0.4043	0.7817	0.7425	-0.0007	0.4092
-0.9385	0.4407	-0.806	0.7633	0.2067	-0.08621
0.8161	0.3082	-0.7845	0.7389	-0.8739	-0.8058
CASE 11					
0.8195	0.9481	-0.8054	-0.7834	0.9682	0.8716
0.9544	0.9542	-0.8141	-0.7414	0.9716	0.8716
0.9535	0.9533	-0.7336	0.2935	0.9548	0.578
0.9594	0.9624	-0.7926	0.6211	0.9749	0.8057
0.4876	0.5018	-0.1323	0.6659	0.9404	0.6712
-0.1635	0.2024	0.5981	0.6964	0.7098	0.7369
-0.9567	0.5915	-0.2627	0.6624	0.9491	0.6978

0.8574	-0.1603	-0.4359	0.7411	-0.8573	0.4807

The simulation results are listed in Table 3. All study cases have the same simulation process but with different configuration. The overall simulation results show that ESN always performs better in terms of prediction accuracy.

Table 3. The simulation results

Case	ESN	TSNN	FFNN
1	4.8e-4	1.5e-3	1.5e-3
2	9.0e-5	3.7e-4	2.3e-3
3	7.8e-4	1.8e-3	6.0e-3
11	9.7e-4	3.0e-3	3.9e-3
12	4.8e-4	1.8e-3	2.2e-3
13	1.9e-3	2.5e-3	5.6e-3

5 CONCLUSION

This paper introduces a new technique for predicting the tool wear states based on statistic analysis and echo-state networks (ESN). First, the statistic analysis is used for selecting valuable features of inputs, then the selected features are fed into ESN to train the neural network and predict subsequent tool- wear states. The simulation results show that our method performs better than TSNN and FFNN.

To the best of our knowledge, this is the first time ESNs were used for milling-tool-wear prediction. As the cutting process is essentially a time series process, so ESNs have a great potential in the predictive-maintenance problem of Cyber-Physical Production Systems (CPPS). Therefore, our future research will focus on adaptive ESNs for CPPS.

ACKNOWLEDGEMENT

This work is funded by the Doctoral College - Cyber-Physical Production Systems (CPPS) Project, Vienna University of Technology, Austria. The authors would like to acknowledge the assistance of Dr. Mohamed Amin Ben Sassi for the excellent suggestions given during the preparation of this paper.

6 REFERENCES

Agogino, A & Goebel, K . 2007. Milling Data Set. In USA: NASA Ames Prognostics Data Repository

Ai, C.S.et al. 2012. The milling tool wear monitoring using the acoustic spectrum. The International Journal of Advanced Manufacturing Technology, 61(5-8), pp.457-463.

Dong,J.etal.,2006.Bayesian-inference-based neural networks for tool wear estimation. The International Journal of Advanced Manufacturing Technology, 30(9-10), pp.797-807.

De Jesus, R.T.R. et al. 2003. Driver current analysis for sensorless tool breakage monitoring of CNC milling machines. International Journal of Machine Tools and Manufacture, 43(15), pp.1529-1534.

García-Nieto, P.J. et al.2015. A new predictive model based on the PSO-optimized support vector machine approach for predicting the milling tool wear from milling runs experimental data. The International Journal of Advanced Manufacturing Technology, pp.1-12.

Goebel, K.F. 1996. Management of uncertainty in sensor validation, sensor fusion, and diagnosis of mechanical systems using soft computing techniques. University of California, Berkeley.

Ghosh, N et al. 2007. Estimation of tool wear during CNC milling using neural network-based sensor fusion. Mechanical Systems and Signal Processing, 21(1), pp.466-479.

Jaeger, H. & Haas, H. 2004. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. science, 304(5667), pp.78-80.

Lukosevicius, M. 2012. A practical guide to applying echo state networks. In Neural networks: Tricks of the trade (pp. 659-686). Springer Berlin Heidelberg.

Mosallam, A et al. 2014. Time series trending for condition assessment and prognostics. Journal of Manufacturing Technology Management, 25(4), pp.550-567.

Palanisamy,P. et al. 2008. Prediction of tool wear using regression and ANN models in end-milling operation. The International Journal of Advanced Manufacturing Technology, 37(1-2), pp.29-41.

Shi, Z. & Han, M. 2007. Support vector echo-state machine for chaotic time-series prediction. IEEE Transactions on Neural Networks, 18(2), pp.359-372.

Wang, J. et al. 2013. Tool life prediction for sustainable manufacturing. 10.14279 /depositonce-3753.