Energy Consumption Data Based Machine Anomaly Detection

Hui Chen, Xiang Fei
Department of Computing
Coventry University
Coventry, UK
{chenh10, x.fei}@coventry.ac.uk

Sheng Wang, Xin Lu, Guoqin Jin, Weidong Li
Department of Mechanical, Automotive and Manufacturing Engineering
Coventry University
Coventry, UK
{ab5337, ab5603, ab4395, weidong.li}@coventry.ac.uk

Xuyang Wu
Department of Computer Science
University College London
London, UK
xuyang.wu.13@ucl.ac.uk

***Abstract*—The ever increasing of product development and the scarcity of the energy resources that those manufacturing activities heavily rely on have made it of great significance the study on how to improve the energy efficiency in manufacturing environment. Energy consumption sensing and collection enables the development of effective solutions to higher energy efficiency. Further, it is found that the data on energy consumption of manufacturing machines also contains the information on the conditions of these machines. In this paper, methods of machine anomaly detection based on energy consumption information are developed and applied to cases on our Syil X4 computer numerical control (CNC) milling machine. Further, given massive amount of energy consumption data from large amount machining tasks, the proposed algorithms are being implemented on a Storm and Hadoop based framework aiming at online real-time machine anomaly detection.**

Keywords—anomaly detection; energy consumption; manufacturing; artificial neural network; Hadoop; Storm

I. Introduction

Product development and manufacturing activities are the major economical pillars of the EU and the world. However the current assumption that natural resources are infinite and that the regenerative capacity of the environment is able to compensate for all human action is no longer acceptable [1]. The scarcity of the conventional energy resources, such as fossil fuel coal, oil or gas, has hindered the development of the global economies [2]. Further, the consumption of these resources may become environmentally hazardous enhancing global warming [3][4]. So how to improve the energy consumption efficiency is of great significance. The sensing technology and wireless sensor network (WSN) technology enables rich and detailed measurements and transmissions of physical phenomenon including energy consumptions of manufacturing equipment. With the sensed data on energy consumption, effective mechanisms for improving the energy efficiency can be developed. Further, it has been found that the energy consumed by machine tools during machining is significantly greater than the theoretical energy required in chip formation, and the data on energy consumption of manufacturing machines also contains the information on the conditions of these machines [5]. Normally the energy consumption anomalies indicate the system is not working at (sub)optimal state or even not functioning, e.g. the performance degradation of facility implying decreasing reliability as well as rising energy consumption [6], the potential failure in spindle due to anomalous power usage spikes [5], and the sudden increased power consumption due to the tool crash detected in our system. The early and effective detection of machine anomaly not only help prevent the breakdown of machine tools and thus reduce the overall production time and cost, but also contribute to higher energy efficiency.

With the advancement of manufacturing models (such as virtual manufacturing, agile manufacturing, etc.) and ICT technologies (such as the Internet, sensor networks and embedded system, especially the cloud computing), cloud manufacturing has become a hot research topic that aims at developing a computing and service-oriented manufacturing model on the Cloud [7][8]. In the meanwhile, more and more data across the product lifecycle are generated that needs to be analysed and mined to support decision making of various levels. One example is the energy consumption data for anomaly detection. With massive production and massive customization, how to effectively and timely detect anomalies from large amount of energy consumption data is challenging. The authors are working towards the implementation of the anomaly detection algorithms on a Storm and Hadoop based framework. Hadoop MapReduce is a software framework for batch processing of data in-parallel on large clusters of commodity hardware in a reliable, fault-tolerant manner [9] while Apache™ Storm is a distributed real-time computation system for processing fast, large streams of data [10].

The rest of the paper is organized as follows. Section 2 introduces the related work. The proposed anomaly detection mechanisms are presented in Section 3. In Section 4, the evaluation of the detection algorithms are provided, followed by the description of the implementation on Storm and Hadoop framework in Section 5. Section 6 concludes our work.

II. Related Work

Anomaly detection, according to [11], is also referred to outlier detection which means the problem of finding patterns in data that do not conform to expected normal behavior. Anomaly detection has been a widely researched problem and finds immense use in a wide variety of application domains. Based on [11], Reference [12] provided a table categorizing the anomaly detection problems as shown in Table 1.

This paper is supported by EU CASES (PIRSES-GA-2011-294931) project and EU Smarter (PEOPLE-2013-IAPP-610675) project.

978-1-4799-8085-7/14 $31.00 © 2014 IEEE
DOI 10.1109/CBD.2014.24

Authorized licensed use limited to: Santa Clara University. Downloaded on March 17,2025 at 04:27:59 UTC from IEEE Xplore. Restrictions apply.

Table 1 Anomaly Detection Categorization

Problem Aspect	Categories
Input data	Binary, categorical, continuous; Univariate, multivariate; Point data, data with structures (time series, sequences, spatial data);
Type of supervision	Supervised, semi-supervised, unsupervised
Anomalies	context-based outliers, instantaneous anomalies, pattern-based outliers, correlation-based outliers

The input data of the machine anomaly detection problem in this paper is univariate, continuous and time series. As the training data is provided only for normal class, the anomaly detection algorithms in this paper are semi-supervised. Further, the anomalies are context based as the input data sample is considered as an anomaly depending on when it occurs; they are also pattern based as they are detected by examining the consecutive occurrences (instantaneous occurrence may be due to noises).

So far a significant number of anomaly detection techniques have been developed. While some of the techniques are generic and can be applied to different application problems, many of them are focused on solving particular types of problems in an application domain[12]. Among those general anomaly detection techniques (such as statistical anomaly detection approaches, classification based approaches and cluster-based approaches) and specific approaches for time series (such as data based approaches and model-based approaches), this paper proposed two approaches, one is Mahalanobis distance based statistical approach and another one is model (artificial neuro network) based approach, to evaluate the effectiveness of various approaches. In addition, application domain knowledge is adopted to interpret and recognize abnormal patterns from those normal ones.

In manufacturing environment, anomaly detection techniques have also been studied to add value to the existing condition based maintenance (CBM) systems. A variety of external sensors, such as accelerometers, acoustic sensors, thermocouples, chip detectors, and stress sensors have been utilized in CBM systems to collect data indicating the performance of mechanical systems[13][14]. Further, reference [15] proposed a strategy to make use of operational data obtained from the machine's controller and signals obtained from external sensors for anomaly analysis/detection within each operating condition. However, none of them incorporates machine energy consumption data in their anomaly detection.

The sensing systems that monitor energy consumption on manufacturing devices have been developed. In [16] a power sensor EDA9033A (Shandong Lichuang Technology Co., Ltd, China) was installed to measure the input power of spindle system to verify the cutting power estimation and display the real-time energy efficiency. Likewise, in [17], a PPC-3 portable power cell was adopted to validate the estimation of mechanical energy requirements of the spindle and feed axes of a machine tool system. In neither case is the energy consumption data collected for anomaly detection.

Both [18] and [19] developed model-based approaches for detecting anomalous energy consumptions in distributed manufacturing systems. Models were learned and the anomalies were detected by the comparison of the predicted and the observed behaviour. The papers didn't explicitly mention whether the anomaly detection is based on instantaneous occurrence or consecutive occurrences as instantaneous occurrence based detection may cause high number of false positive due to the signal noises

. Also they didn't take into consideration how to provide online real-time services given the ever increased energy consumption data set.

To the extent of our knowledge, we summarize the contributions of this paper as follows:

- raw energy consumption data from a real machine is collected and analyzed; also domain knowledge and pattern based approach are used to improve the accuracy of the anomaly detection.
- a prototype is being implemented that integrates our proposed approaches onto a Storm and Hadoop framework to provide online real-time anomaly detection services via massive energy consumption data sets.

III. Anomaly Detection Algorithms

As mentioned in section II, the input data for this paper is the energy consumption data that is univariate, continuous and time series. In the future, other types of sensor data can be incorporated if necessary. The anomaly detection algorithms are semi-supervised as it is hard to enumerate and collect all the possible anomalies. Further, the anomalies are both context based and pattern based to improve the accuracy and reduce the false alarms.

In this paper, two anomaly detection algorithms were designed, one is Mahalanobis distance based statistical approach and another one is artificial neural network (ANN) based approach.

In order to provide formal description of our algorithms, based on the definitions provided by [20], below are the definitions used in this paper

Definition 1: (Temporal Sequence)

A temporal sequence, or a time series, is represented as $X(t)$, where $t = 0, \ldots, N-1$. In this paper, $X(t)$ represents energy consumption data, and can be extended to other time series such as vibration data, operational data, etc. Note that $X(t)$ is a stochastic process as each element, $x(t)$, in the sequence is a random number.

Definition 2: (Model)

A model, demoted by $\mathrm{M}_{\mathrm{X}}(t)$ to represent our knowledge about the underlying temporal sequence up to t. This model can be a physics-based model provided by domain experts, or a

Authorized licensed use limited to: Santa Clara University. Downloaded on March 17,2025 at 04:27:59 UTC from IEEE Xplore. Restrictions apply.

model constructed from available data x(t) where $t = 1, \cdots, t_0$, e.g. an ANN model.

Definition 3: (Matching Value and Matching Function)

The matching function, denoted as $F(\mathrm{M_X}(t_0 - 1), \mathrm{x}(t_0))$, is a function that can quantify how well the model $\mathrm{M_X}(t_0)$ matches the temporal sequence. The match value $V(t_0) \in R$, is defined as $V(t_0) = F(\mathrm{M_X}(t_0 - 1), x(t_0))$. In this paper, we use ANN to predict $\tilde{x}(t_0) = \mathrm{M_X}(t_0 - 1)$ based on $\mathrm{X}(t_0 - 1)$, and compare it with the sensed data $\mathrm{x}(t_0)$.

Definition 4: (Occurrence)

Denoted by $O(t_0)$, occurrence at t_0 is defined as

$$O(t_0) \equiv I\{D(t_0) \notin (\varepsilon_1(t_0), \varepsilon_2(t_0))\} \qquad (1)$$

where $I\{\cdot\}$ is the indicator function, and $\varepsilon_2(t_0) - \varepsilon_1(t_0) > 0$ is a predefined tolerance width.

For different types of detection approaches, $D(t_0)$, the detection function, is different. For model based anomaly detection, $D(t_0) = V(t_0)$ and

$$O(t_0) \equiv I\{V(t_0) \notin (\varepsilon_1(t_0), \varepsilon_2(t_0))\} \qquad (2)$$

Definition 5: (Surprise)

A surprise is observed if $O(t_0) = 1$

Definition 6: (Event and Event Duration)

Denoted by $E_n(t_0)$, an event at time t_0 is defined as

$$E_m(t_0) = [O(t_0)\ O(t_0+1)\ \cdots\ O(t_0+m-1)]^{T} \qquad (3)$$

where the $1 - norm$ of $E_n(t_0)$, denoted as $|E_n(t_0)|$, is

$$|E_n(t_0)| = \sum_{i=0}^{n-1} O(t_0 + i) = m \qquad (4)$$

$|E_n(t_0)|$ is the number of surprises happened consecutively in the event $E_n(t_0)$.

Definition 7: Anomaly Event

Event $e_n(t_0)$ is defined as an anomaly event if $e_n(t_0)$ satisfies

$$|e_n(t_0)| > h \qquad (5)$$

where h is a predefined threshold with $h \in N$.

In this paper, two anomaly detection algorithms were designed. They share the same steps as shown in Table 2. They differ in step 3 and step 4.

Table 2 Steps of anomaly detection

Steps	description
1	Determine a manufacturing process.
2	Collect the energy consumption data during the normal (anomaly free) operation.
3	Analyze, or build models from the collected data
4	With the results from step 3, detect the anomalies given the energy consumption data from the real processes.

For any new anomaly detection algorithm, what need to determine are $D(t)$, $\varepsilon_1(t_0)$, $\varepsilon_2(t_0)$ and h.

A. Mahalanobis Distance based Statistical Approach

Definition 8: (Mahalanobis Distance)
Mahalanobis distance is defined as

$$D^{mahal}(X) = \sqrt{(X - \vec{\mu})^T COV^{-1}(X - \vec{\mu})} \qquad (6)$$

As the input data for this paper is univariate which is the special case of multivariate data, the $D^{mahal}(X)$ for the anomaly detection in this paper can be simplified as

$$D^{mahal}(x) = \sqrt{(x - \mu)^2 VAR^{-1}} \qquad (7)$$

Mahalanobis distance based approach is extensible in that it can accommodate new types of sensor data or operational data when in the future we need to incorporate them into our anomaly detection approach.

Given the formula (7), the step 3 of Mahalanobis distance based approach can be elaborated in Table 3, and step4 in Table 4.

In Table 4, $\varepsilon_1(t_0)$ is set to zero as $D^{mahal}(x(t)) \geq 0$. $\varepsilon_2(t_0)$ is set to 3 following *3-sigma rule*. This is based on the assumption that the energy consumption, when the machine is in working (in-cycle) state, follows Gaussian distribution. h is set to 5 to filter out the possible high frequency noises that cause $D^{mahal}(x) > \varepsilon_2$ instantaneously.

Table 3: Mahalanobis distance based data analysis

Step 3	description
1	Calculate the mean of the data set obtained from Step 2 $\mu = E[\mathrm{X}(t)] = \sum_{i=0}^{N-1} \mathrm{x}(i)$
2	Calculate the standard deviation of the data set obtained from Step 2 $\sigma = \sqrt{E[(\mathrm{X}(t) - \mu)^2]}$

Table 4: Mahalanobis distance based anomaly detection

Step 4	description
1	Set $D_x(t)$ in (1) as $D^{mahal}(\mathrm{x}(t))$
2	Set $\varepsilon_1(t_0) = 0$
3	Set $\varepsilon_2(t_0) = 3$
4	Set h in (5) = 5
5	For each energy consumption data from the manufacturing processes
6	Calculate $D^{mahal}(\mathrm{x}(t))$
7	Check whether the condition for an anomaly event defined in (5) holds
8	If anomaly event is detected, send an alarm

Authorized licensed use limited to: Santa Clara University. Downloaded on March 17,2025 at 04:27:59 UTC from IEEE Xplore. Restrictions apply.

Table 5: ANN based learning and modelling

Step 3	description
1	Set $M_X(t)$ as a 10-input and 1-output ANN
2	Set the input as $[(i-6),\ (i-5)\ ,\ \cdots\ ,(i-1)\ ,\mathrm{x}(i-6),\ \ \mathrm{x}(i-5)\ ,\cdots,\ \mathrm{x}(i-1)]$ and output $\mathrm{x}(i)$
3	Given the data set obtained from Step 2, construct a 10-input, 7-hidden and 1-output ANN and learn the parameters(weights), with the aid of RapidMiner

Table 6: ANN based anomaly detection

Step 4	description
1	Set $D(t)$=$V(t)$= $F(\mathrm{M_X}(t-1), x(t))$ = $\lvert \mathrm{M_X}(t-1) - x(t)\rvert / \mathrm{M_X}(t-1)$
2	Set $\varepsilon_1(t_0) = 0$
3	Set $\varepsilon_2(t_0) = 0.1$
4	Set h in (5) = 5
5	For each energy consumption data from the manufacturing processes
6	Calculate $V(t)$
7	Check whether the condition for an anomaly event defined in (5) holds
8	If anomaly event is detected, send an alarm

B. Model (ANN) based Approach

Among those models that can be used to model the energy consumption data, such as linear regression, support vector regression, etc., ANN based approach was adopted due to its robustness on noisy training data and high speed to evaluate new examples, especially due to its massively parallel processing capability that makes it easier to be integrated to Hadoop framework [21]. During the design and simulation phase, RapidMiner [22], an integrated environment for machine learning and data mining, was used and the step 3 and step 4 listed in Table 2 were expanded as shown in Table 5 and Table 6 respectively.

The energy consumption model is learned by means of a multi-layer perceptron with a feed-forward neural network trained by a back propagation algorithm. The sigmoid function is applied as an activation function for each neuron.

The input of the ANN contains not only 5 previous energy consumption data $[\mathrm{x}(i-6),\ \ \mathrm{x}(i-5)\ ,\cdots,\mathrm{x}(i-1)]$, but also the timestamps $[(i-6),\ (i-5)\ ,\ \cdots,(i-1)\ ]$. This is because the anomalies are context based and as will be seen in Figure 8, without considering the timestamps, ANN will not be able to tell the anomalies from the normal data.

Also, as the timestamps and the energy consumption data are in different ranges, normalization is an essential step before applying those data to the ANN.

IV. CASE STUDY

The energy consumption of a facemilling process using a Syil X4 CNC milling machine was monitored using our in-house WSN system (Due to space limitations, our energy consumption WSN system will be described in another paper). According to [5], the operational states of the machine tool include *startup*, *shutdown*, *idle*, and *in-cycle* (machining a part), and there exist ''spikes'' in the energy consumption during a spindle speed change. This domain specific knowledge makes the anomaly detection non-trivial as we need to tell from the expected spikes the unexpected spikes that imply anomalies. In the meanwhile, this knowledge help improve the accuracy of the detection. The WSN system senses and collects the overall current that drives the face milling of an aluminum alloy. The CNC machines are controlled by G-code programmed to perform specific tasks.

The first task is to cut a face of 0.5mm with the spindle speed of 1800 rpm. The raw data is shown in Figure 1, from which we can see that

- An energy spike appears at the beginning of the milling operation, which conforms to the expected spikes mentioned in [5].
- During the in-cycle state, the data is fluctuated due to both process and measurement high-frequency noises. As a result, the sensed the data $\mathrm{x}(t)$ is taken as random number. The fluctuation makes the anomaly detection challenging as the noises will cause false-alarms and we need to filter out those noises from the real anomalies.

A. Low-pass filtering of Raw Data

In order to lessen the impact of noises on the data quality, the following measures were taken

- Multiple independent normal operations $\mathrm{X}_i(t), i \in [0,\cdots,M-1]$ were conducted (in this case M is 7, and more will be carried out in the future), and the aggregated $\overline{\mathrm{X}(t)} = \frac{1}{N}\sum_{i=0}^{M-1}\mathrm{X}_i(t)$ is the data set for training. According to central limit theorem (CLT), as the energy consumption data, process noise and measurement noise are independent, each $\overline{\mathrm{x}(t)}$ tends to follow normal distribution.
- The moving average (MA) filter is adopted as the low-path filter applied to $\overline{X(t)}$. In spite of its simplicity, the MA filter is optimal for a common task such as reducing random noise [23]. The MV filtering is triggered after the expected *startup* spike.

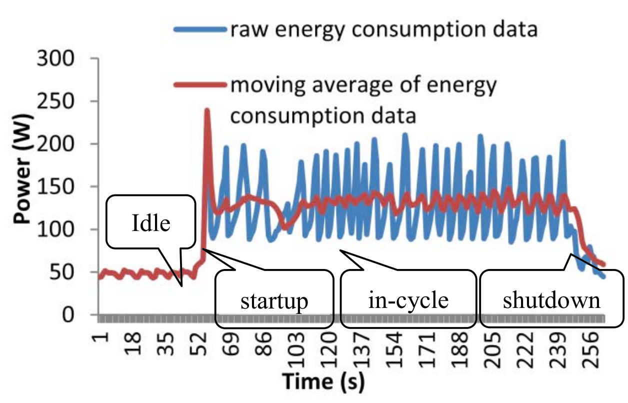

Figure 1: raw and MA of energy consumption data

Authorized licensed use limited to: Santa Clara University. Downloaded on March 17,2025 at 04:27:59 UTC from IEEE Xplore. Restrictions apply.

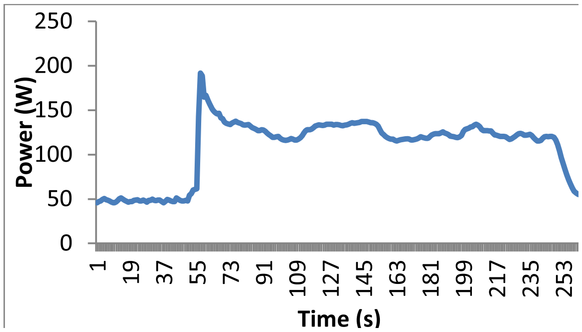

Figure 2: MA of aggregated normal data (1800 rpm)

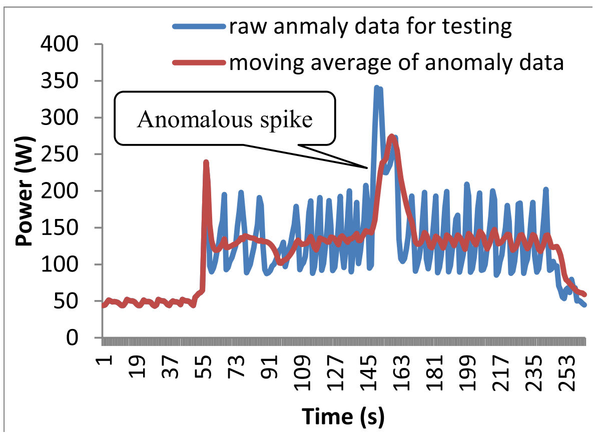

Figure 3: Anomaly data set (1800 rpm)

Figure 2 shows the moving average of the aggregated energy consumption data from normal operations.

B. Synthesis of Anomaly Data Set

According to [5] and our experience of a machine tool crash, we synthesize an anomaly data set, as shown in Figure 3, for testing the anomaly detection approaches.

C. Mahalanobis distance based statistical approach

The initial *startup* spike triggers the distance based algorithm. Each sensed data is processed following Table 2, Table 3 and Table 4. Figure 4 illustrates the anomaly detection result where it can be see that

- For a normal operation, the Mahalanobis distance is less than the tolerance width $(\varepsilon_2(t_0) - \varepsilon_1(t_0))$ which is 3 in this case.
- An anomaly starts at time 83, and the first time when the distance is greater 3 is at time 88 and the anomaly is detected at time 93 (h is set to 5). The delay in the detection results from the setting of $\varepsilon_1(t)$ and $\varepsilon_2(t)$, the setting of h, and the use of moving average. Reduce the tolerance width will speed up the detection but will also increase the possibility of false alarms.

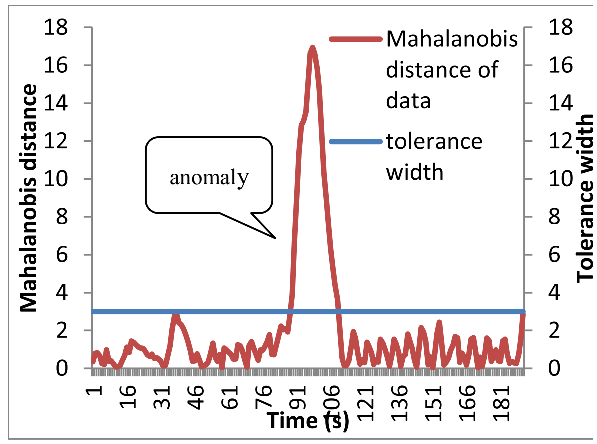

Figure 4: Mahalanobis distance of anomaly data set

D. ANN based approach

To further study the anomaly detection during more complicated processes, in second scenario, the spindle speed changes from 1800 rpm to 3600 rpm in the middle of the face milling process and the synthesized energy consumption data for normal processes is indicated in Figure 5. An unexpected spike occurs before the spindle increases its speed to 3600 rpm as shown in Figure 6. This makes the anomaly detection challenging as the algorithm may take the unexpected spike as the spindle speed increase.

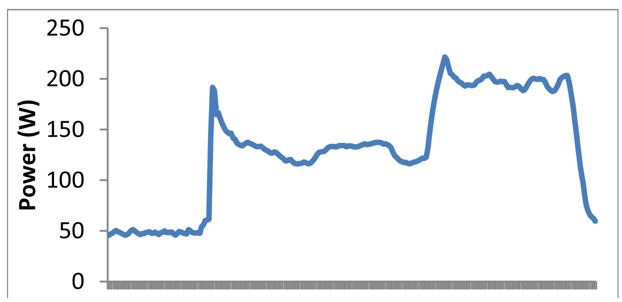

Figure 5: MA of aggregated normal data (1800 and 3600 rpm)

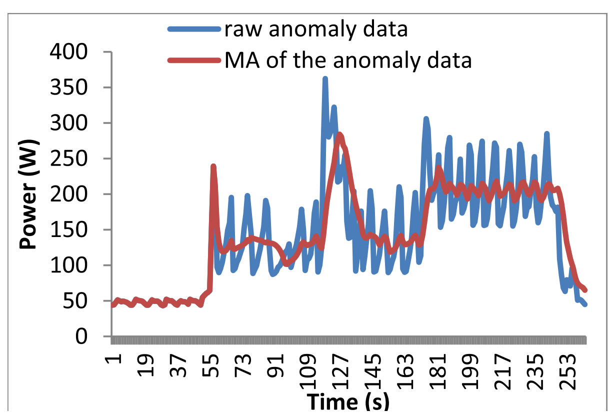

Figure 6: Anomaly data set (1800 and 3600 rpm)

Authorized licensed use limited to: Santa Clara University. Downloaded on March 17,2025 at 04:27:59 UTC from IEEE Xplore. Restrictions apply.

A 10-inout, 7-hidden and 1-output ANN is constructed and trained from the aggregated normal data set ($\overline{X(t)}$) following the procedures in Table 5. 10-fold cross validation is performed and the root mean squared error (RMSE) for the model is 5.74. The anomaly detection process is listed in Table 6 and triggered by the startup spike. Figure 7 demonstrates the detection result of the ANN based anomaly detection, where it can be observed that

- An anomaly starts at time 63, and the first time when the tolerance width is greater 0.1 is at time 67 and the anomaly is detected at time 72 (h is set to 5). The delay in the detection results from the setting of the tolerance width ($\varepsilon_2(t_0) - \varepsilon_1(t_0)$), the setting of h, and the use of moving average.
- During the whole in-cycle process, some spikes of matching value that is greater than 0.1 also appear. This may be due to the noises incurred and the learning accuracy of ANN. However, as their duration is less than h which is set as 5, they are not taken as anomalies. So the proposed pattern based detection is able to filter out the instantaneous anomalous occurrences, and reduce the possibility of false alarms.
- Figure 8 illustrates the result of ANN based anomaly detection without taking timestamps as the inputs of ANN. It can be observed that if the timestamps are not taken into account during the training stage, the ANN will not be able to tell the anomaly as the anomalous spike is taken as an expected spike for the spindle speed increase.

E. Comparison between Mahalanobis distance based approach ANN based approach

For Mahalanobis distance based approach, the computational complexity for anomaly detection is $O(n^2)$ where n is the dimension of X in (6) and in our case $n = 1$; while for ANN based approach, the complexity is $O(i \times h) + O(h \times o)$ where i, h, and o are the number of neurons in input layer, hidden layer and output layer respectively and in our case $i = 10$, $h = 7$ and $o = 1$. So compared to ANN based approach, Mahalanobis distance based approach is simpler in computation in our case.

On the other hand, for Mahalanobis distance based approach, to be able to detect the anomalies for processes with multiple spindle speed settings, Mahalanobis distance should be calculated separately for each spindle speed. Further information, such as the timestamps, is needed to tell whether the spike detected is expected or not. All these make the Mahalanobis distance based approach more complex compared ANN based approach under the assumption that the ANN is well designed and trained.

In terms of the performance of the anomaly detection, for both Mahalanobis distance based approach and the ANN based one, the time to detect the anomaly depends on the setting of tolerance width and the setting of h. Compared to Fighre 7, the anomaly using Mahalanobis distance as shown shown in Figure 4 is more distictive, which leads to the lower possibility of false alarms.

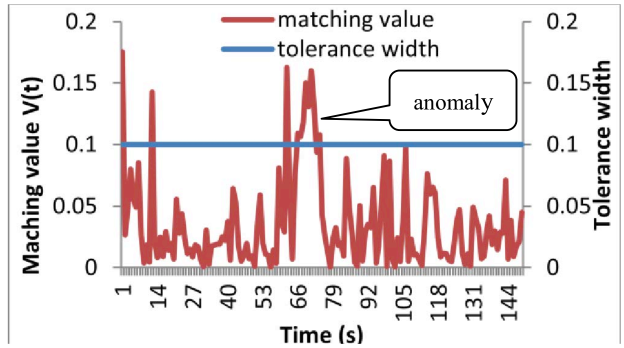

Figure 7: Result of ANN based anomaly detection

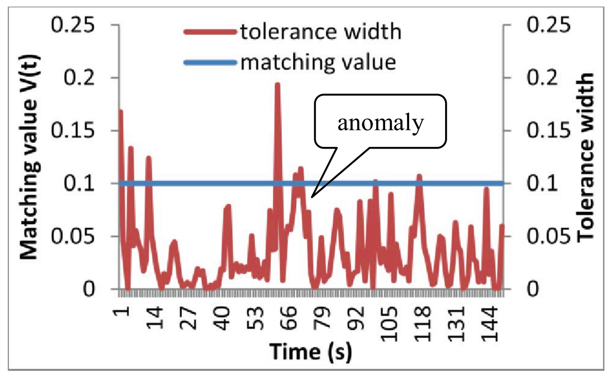

Figure 8: Result of ANN based anomaly detection without taking timestamps as inputs of ANN

V. TOWARDS STORM AND HADOOP BASED REAL-TIME ANOMALY DETECTION

To provide real-time anomaly detection service via analyzing potentially massive energy consumption data, Storm and Hadoop MapReduce frameworks, as shown in Figure 9, are adopted where the off-line normal behavior learning and modelling are implemented on Hadoop MapReduce while on-line anomaly detection on Storm. The energy consumption WSN, Storm and Hadoop are interacted either synchronously via Sockets or asynchronously via Apache Kafka. So far the core functions such as statistical calculations such as mean, variance, and correlation, etc. have been implemented on Hadoop Mapreduce, the unit testing on the ANN algorithm is being carried out. The system testing on real machines in the manufacturing environment will be conducted as our future work.

Authorized licensed use limited to: Santa Clara University. Downloaded on March 17,2025 at 04:27:59 UTC from IEEE Xplore. Restrictions apply.

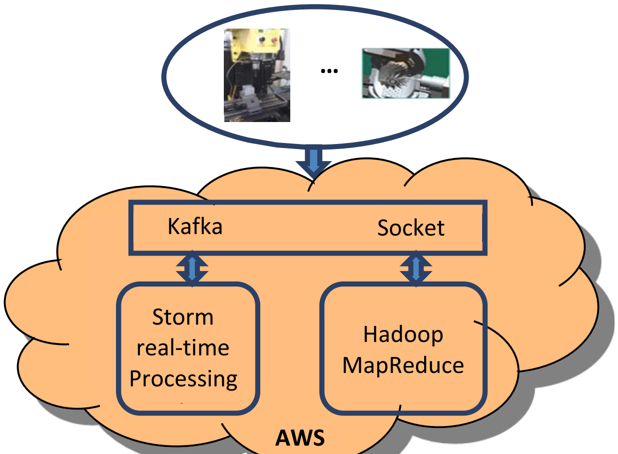

Figure 9: Structure of Storm and Hadoop based real-time anomaly detection system

VI. Conclusions

In this paper, two approaches to detecting machine anomalies via energy consumption data are proposed. One approach is Mahalanobis distance based statistical approach and another one is model (ANN) based approach. Real and synthesized energy consumption data is used for analyzing, training, modelling and detecting anomalies. For complex production processes, well designed and trained model (such as ANN) proves to make the detection task more straightforward compared to statistical approaches although Mahalanobis distance has lower computational complexity for low dimensional vectors. Domain knowledge and pattern based approach are adopted to improve the accuracy of the anomaly detection. Further, implementation of the proposed approaches on Storm and Hadoop framework is been carried out aiming at real-time detection via massive energy consumption data sets. Apart from carrying on the system implantation, other types of data such as vibration and temperature will be incorporated and new approaches will also be studied to improve the effectiveness and performance of the machine anomaly detection.

References

[1] M. Garetti, and M. Taisch, "Sustainable manufacturing: trends and research challenges," Production Planning and Control, vol. 23 (2-3), February-March, pp. 83-104, 2012.

[2] F. Jovane, H. Yoshikawa, L. Alting, C. R. Boer, E. Westkamper, D. Williams, et al. "The incoming global technological and industrial revolution towards competitive sustainable manufacturing," CIRP Annals – Manufacturing Technology, vol. 75, pp. 641-659, 2008.

[3] S. Bilgen, "Structure and environmental impact of global energy consumption," Renewable and Sustainable Energy Reviews, vol. 38, pp. 890–902, 2014.

[4] G.Q. Jin and W.D. Li, "Life cycle management of LCD televisions – A case study," Handbook of Manufacturing Engineering and Technology, 2014, Book chapter, Springer.

[5] A. Vijayaraghavan and D. Dornfeld, "Automated energy monitoring of machine tools," CIRP Annals - Manufacturing Technology, vol. 59 pp. 21–24, 2010.

[6] J. Yan and D. Hua, "Energy consumption modeling for machine tools after preventive maintenance," IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 2201 – 2205, Dec. 2010.

[7] W.D. Li,., K. Xia, B. Lu, K.M. Chao, L. Gao, and J.X. Yang, "A distributed service of selective disassembly planning for waste electrical and electronic equipment with case studies on liquid crystal display," Springer Series in Advance Manufacturing: Cloud Manufacturing 2013

[8] F. Tao, L. Zhang, V. C. Venkatesh, Y. Luo, and Y. Cheng, "Cloud manufacturing: a computing and service-oriented manufacturing model," J. Eng. Manuf. Vol. 225, pp. 1969–1976, 2011.

[9] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters," Communications of the ACM , vol. 51 (1), pp. 107-113, January 2008

[10] Apache Incubator, "Storm, distributed and fault-tolerant realtime computation," https://storm.incubator.apache.org/

[11] V. Chandola, A. Banerjee, and V. Kumar, (2009). "Outlier detection: A survey,". ACM Computing Surveys, vol. 41 (3), pp. 1– 72, July 2009.

[12] L. Li, "Anomaly detection in airline routine operations using flight data recorder data," http://hdl.handle.net/1721.1/82498

[13] D. Djurdjanovic, J. Lee, and J. Ni, "Watchdog Agent—an infotronics-based prognostics approach for product performance degradation assessment and prediction," Advanced Engineering Informatics , vol. 17 (3-4), pp. 109-125, 2003..

[14] R. Kothamasu, S. H. Huang, and W. H. VerDuin, "System health monitoring and prognostics – a review of current paradigms and practices." International Journal of Advanced Manufacturing Technology, vol. 28, pp. 1012-1024, 2006.

[15] L. Liao and P. Pavel, "Machine anomaly detection and diagnosis incorporating operational data applied to feed axis health monitoring," International Manufacturing Science and Engineering Conference (MSEC), 2011

[16] S. Hu, F. Liu, Y.n He, and T. Hu, "An on-line approach for energy efficiency monitoring of machine tools," Journal of Cleaner Production, vol. 27 pp.133 – 140, 2012.

[17] O. I. Avram and P. Xirouchakis, "Evaluating the use phase energy requirements of a machine tool system," Journal of Cleaner Production vol. 19, pp. 699-711, 2011.

[18] S. Faltinski, H. Flatt, F. Pethig, B. Kroll, A. Vodencarevic, A.Maier, et al. "Detecting anomalous energy consumptions in distributed manufacturing systems," 10th IEEE International Conference on Industrial Informatics (INDIN), pp. 358 – 363, July 2012.

[19] S. Windmann, S. Jiao, O. Niggemann, H. Borcherding, "A stochastic method for the detection of anomalous energy consumption in hybrid industrial systems," 11th IEEE International Conference on Industrial Informatics (INDIN), pp. 194 – 199, July 2013.

[20] J. Ma and S. Perkins, "Online novelty detection on temporal sequences," Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 613-618, 2003

[21] Z. Liu, H. Li, and G. Miao, "MapReduce-based Backpropagation Neural Network over Large Scale Mobile Data," Sixth International Conference on Natural Computation (ICNC), pp. 1726 – 1730, Aug. 2010.

[22] Rapidminer, http://rapidminer.com/

[23] H. Azami, K. Mohammadi, and B. Bozorgtabar, "An improved signal segmentation using moving average and savitzky-golay filter," Journal of Signal and Information Processing, vol. 3, pp. 39-44, 2012.

[24] X.Lu and W. Li, "A systematic review on industrial wireless sensor networks," to appear in the Proceeding of Sustainable Design and Manufacturing, 2014

Authorized licensed use limited to: Santa Clara University. Downloaded on March 17,2025 at 04:27:59 UTC from IEEE Xplore. Restrictions apply.