ARTICLE IN PRESS

ISA Transactions xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

ISA Transactions

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

Practice article

# Hybrid algorithmic approach oriented to incipient rotor fault diagnosis on induction motors

Ignacio Martin-Diaz[a,b], Daniel Morinigo-Sotelo[a], Oscar Duque-Perez[a], Roque A. Osornio-Rios[c], Rene J. Romero-Troncoso[c,*]

[a] Elect. Eng. Dept. Escuela de Ingenierias Industriales, Sede Paseo del Cauce, University of Valladolid, Paseo del Cauce, 59, 47011, Valladolid, Spain
[b] HSPdigital-CATelematica, DICIS, Universidad de Guanajuato, Carr. Salamanca-Valle km 3.5 + 1.8, Palo Blanco, 36885, Salamanca, Guanajuato, Mexico
[c] Universidad Autónoma de Querétaro, HSPdigital CA-Mecatronica, Facultad de Ingenieria, Río Moctezuma 249, San Juan del Rio, 76806, Querétaro, Mexico

A R T I C L E I N F O

*Keywords:*
Induction motor
Fault diagnosis
Artificial intelligence
Simulated annealing algorithm
Oblique random forests

A B S T R A C T

This paper investigates the current monitoring for effective fault diagnosis in induction motor (IM) by using random forest (RF) algorithms. A rotor bar breakage of IM does not derive in a catastrophic fault but its timely detection can avoid catastrophic consequences in the stator or prevent malfunctioning of those applications in which this sort of fault is the primary concern. Current-based fault signatures depend enormously on the IM power source and in the load connected to the motor. Hence, homogeneous sets of current signals were acquired through multiple experiments at particular loading torques and IM feedings from an experimental test bench in which incipient rotor severities were considered. Understanding the importance of each fault signature in relation to its diagnosis performance is an interesting matter. To this end, we propose a hybrid approach based on Simulated Annealing algorithm to conduct a global search over the computed feature set for feature selection purposes, which reduce the computational requirements of the diagnosis tool. Then, a novel Oblique RF classifier is used to build multivariate trees, which explicitly learn optimal split directions at internal nodes through penalized Ridge regression. This algorithm has been compared with other state-of-the-art classifiers through careful evaluation of performance measures not encountered in this field.

## 1. Introduction

Induction motors (IMs) predominate in industry due to their low manufacturing costs, power-weight ratio and robustness [1,2]. Copper rotor failures can develop into catastrophic breaking due to bar bending or material projection. Early diagnosis of these faults is a primary concern and will guarantee the operation of motors, safeguarding their integrity [3,4].

Rotor fault detection can be achieved monitoring different signals [2]: vibration [5,7], sound [7], acoustic emission [2], temperature [8], air gap magnetic flux [9], instantaneous power [10,11], supply voltage [3] and stator current [1,12]. However, motor current signature analysis (MCSA) has been preferred [13,14] because it is non-invasive, low-cost, and easy to measure. Besides, the motor shutdown is not required.

Many works have proposed fault identification techniques based on the analysis of the stator current. Authors in Ref. [15] improved the Park's Vector (PV) approach for detecting a broken rotor bar (BRB). They tracked the higher harmonic index after the application of elliptic and notch filters on the PV components. In Ref. [16], the authors analysed the zero sequence current (ZSC) with MUSIC to increase the BRB detection reliability. Authors in Ref. [17] tested the ZSC spectrum for detecting rotor asymmetries. The results are promising, but the motor must have a delta connection or the neutral-connected. Although it requires three current sensors, it shows advantages that complement other methods. Samanta et al. [18] used an extended-Kalman based signal conditioner to remove the fundamental component of the stator current signal. The authors claim that their system is fast, accurate and can be implemented online. Bessam et al. [19] presented a NN approach where the Hilbert Transform is used for the diagnosis of BRBs at low load. Authors in Ref. [20] proposed an analytical equation that relates

*Abbreviations:* IM, induction motor; CBM, condition based monitoring; MCSA, motor current signature analysis; BRB, broken rotor bar; ANN, artificial neural network; DT, decision tree; PCA, principal component analysis; RF, random forests; SA, simulated annealing; ORF, oblique RF; KNN, k-Nearest Neighbours; FFT, fast Fourier transform; PM, performance metric; AUC, area under an ROC curve; ROC, receiver operating characteristic; OOB, out of bag; OVA, one versus all; TPR, true positive rate; FPR, false positive rate

* Corresponding author.
*E-mail address:* troncoso@hspdigital.org (R.J. Romero-Troncoso).

https://doi.org/10.1016/j.isatra.2018.07.033
Received 6 November 2017; Received in revised form 3 April 2018; Accepted 24 July 2018


Please cite this article as: Martin-Diaz, I., ISA Transactions (2018), https://doi.org/10.1016/j.isatra.2018.07.033

the number of BRB to the IM stator vibrations. Abd-el-Malek et al. [21] presented a technique based on the statistical analysis of the stator current envelope for detecting the exact fault location. A simulation under two operating modes validates the method, which otherwise lacks experimental verification. Interesting improvements of actual techniques were recently published, but the diagnosis of incipient faults based on statistical learning is not yet studied. New features, feature selection methods, classifier types, tuning and classi- cation methods continue emerging to enhance systems performance and diagnostic accuracy [12,22]. Thus, as suggested in Ref. [22] improvement of existing diagnostic methods and discovery of new fault indicators is a necessity. Sideband harmonics around the fundamental component of the stator current are the most used fault features [23]. Gyftakis et al. [17,24] propose thresholds for the fault detectability in line-fed IM. However, diagnosis of inverter-fed IM remains to be a topic of interest in the fault diagnosis community. Bruzzese [25] proposes new patterns for IM fed by non-sinusoidal power supplies that demonstrated to detect one or more BRBs.

Recent developments in machine learning algorithms permit to address the automatic diagnosis of IM faults [26]. The standard methodology consists of the following steps: (1) a collection of suited measurements related to the fault of interest; (2) extraction of fault signatures or features; (3) application of a feature reduction or feature selection technique; (4) construction of a classification model. In Ref. [27], the diagnosis is based on the information provided by the discrete wavelet transform. In Ref. [6], the authors used genetic algorithms to select the most significant features and optimize the artificial neural network (ANN) parameters. In Ref. [28], the diagnosis is performed by a multilayer perceptron ANN with statistical parameters as the inputs, whose dimensionality was reduced by Principal Component Analysis (PCA). Authors in Ref. [1] presented an intelligent system based on a combination of stationary wavelet packet transform and multiclass wavelet support vector machines. Researchers in Ref. [29] used an adaptive neuro-fuzzy inference system in combination with decision trees (DT), which permit to build explanatory rules to justify the predictions [30]. Although DT have a low bias, they usually suffer from high variance, which can be solved by combining the predictions of several randomized trees into a single model known as Random Forest (RF). The studies mentioned above suppose an important contribution, and our purpose is to improve some steps of the diagnosis methodology. Our work considers incipient rotor severities in IM fed from the line and inverter. The latter is an important point because this type of feeding introduces noise in the fault signatures, complicating the detection and the diagnosis.

In this paper, we collect and use all fault indicators presented in the recent literature, and for the fault classification, we propose a hybrid approach using the Simulated Annealing (SA) algorithm and the RF classifier. We use a promising version of the RF, known as Oblique RF (ORF). This version uses an oblique node splitting criteria (multivariate decision trees) via Ridge regression that allows improving the classification performance in those cases where the features are correlated. In machine learning, the problem of over fitting is always present when the data available is small. For this reason, we evaluate the effectiveness of the proposed approach with up to 7 experiments, where additional metrics to those used in the literature (Sensitivity, Specificity, ROC) have been considered.

The major contributions of this article are threefold:

- We use all the fault indicators proposed in the recent literature. No previous works have used the indicators proposed in Ref. [25] for the diagnosis of incipient BRB. These indicators were successful at detecting 1, 2, 3 and four broken bars for non-sinusoidal IM supplies.
- We propose the use of the SA algorithm to identify the features with greater discriminant capacity in different experiments. No study has evaluated the discrimination ability of fault features and if that capacity is kept when the IM supply is changed. Unlike genetic algorithms, the optimum design procedures based on the SA are less time consuming, and the optimum solutions obtained may avoid local ones, being feasible both mathematically and practically.
- We use the RF and ORF to classify incipient rotor faults. Their random structure increases their generalization ability. ORF, through its oblique splitting, improves the classification performance in some experiments.

The parts that constitute the present work are organized as follows. Section 2 introduces the theory that justifies the fault signatures suitability. Then, Section 3 describes the algorithms that compose the proposed hybrid approach for diagnosing rotor faults. Experimental results are presented in Section 4 for a line-fed and inverter-fed IM under two different load levels. In this section, a comprehensive comparison between RF, ORF, classification and regression trees (CART) and KNN is shown. The purposed approach is compared with several state-of-the-art techniques found in the literature. Finally, Section 5 concludes this study.

## 2. Related work

This section presents the theory behind the fault signatures extraction to justify their suitability. In this work, the first set of features has two types of data calculated from the measured stator currents: signatures computed from the raw stator current in the time domain, and signatures obtained by spectral analysis of the same signals.

### 2.1. Time-domain fault signatures

A statistical analysis of the raw data (stationary period of the stator current) permitted to calculate fifteen fault-signatures in the time-domain. These signatures are described in Table 1.

### 2.2. Spectral fault signatures

As it is well known, a BRB fault occurs after the development of small cracks at the junction between the bar and the end ring [31]. Consequently, the resultant signatures for cracked and BRBs on the current spectrum are the effects due to rotor circuit asymmetries [32]

**Table 1**
Statistical features extracted from the stator current in time domain.

| Time-domain features | | |
|---|---|---|
| Feature | Symbol | Expression |
| First Moment | $m_1$ | $\frac{1}{N}\sum x[n]$ |
| Second Moment | $m_2$ | $\frac{1}{N}\sum (x[n]-\bar{x}[n])^2$ |
| Third Moment | $m_3$ | $\frac{1}{N}\sum (x[n]-\bar{x}[n])^3$ |
| 4th Moment | $m_4$ | $\frac{1}{N}\sum (x[n]-\bar{x}[n])^4$ |
| Second Cumulant | $c_2$ | $m_2 - m_1^2$ |
| Third Cumulant | $c_3$ | $m_3 - 3m_1m_2 + 2m_1^3$ |
| 4th Cumulant | $c_4$ | $m_4 + 3m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4$ |
| Skewness | *Skew* | $\frac{m_3}{(\sqrt{m_2})^3}$ |
| Kurtosis | *Kurt* | $\frac{m_4}{(\sqrt{m_2})^4}$ |
| Absolute mean | AM | $\lvert\bar{x}\rvert$ |
| Peak value | PV | $\frac{1}{2}(max(x[n]) - min(x[n]))$ |
| Squared root value | SRV | $\left(\frac{1}{N}\sum\sqrt{\lvert x\rvert}\right)^2$ |
| Crest factor | CF | ${}^{PV}/_{RMS}$ |
| Shape factor | SF | ${}^{RMS}/_{\lvert\bar{x}\rvert}$ |
| RMS value | RMS | $\sqrt{\frac{1}{N}\sum_{n=0}^{N-1}[x[n]-\bar{x}]^2}$ |