A Multi-label Classification Approach for the Detection of Broken Bars and Mixed Eccentricity Faults Using the Start-up Transient

George Georgoulas[1], Vicente Climente-Alarcon[2], Jose A. Antonino-Daviu[3], Chrysostomos D. Stylios[4], Antero Arkkio[2], and George Nikolakopoulos[1]

[1]*Control Engineering Group, Department of Computer Science, Electrical and Space Engineering, University of Technology, SE-97187 Luleå, Sweden*

[2]*Department of Electrical Engineering and Automation, Aalto University, P. O. Box 13000, FI-00076 Aalto, Finland*

[3]*Instituto Tecnologico de la Energia, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain*

[4]*Laboratory of Knowledge and Intelligent Computing, Department of Computer Engineering, TEI of Epirus, 47100 Artas, Kostakioi, Greece*

Abstract—**In this article a data driven approach for the classification of simultaneously occurring faults in an induction motor is presented. The problem is treated as a multi-label classification problem with each label corresponding to one specific fault, using the power-set approach. The faulty conditions examined, include the existence of a broken bar fault and the presence of mixed eccentricity with various degrees of static and dynamic eccentricity. For the feature extraction stage, the time-frequency representation, resulting from the application of the short time Fourier transform of the start-up current is exploited. The proposed approach is validated using simulation data with promising results.**

Keywords—*multi-label classification, rotor broken bars, mixed eccentricity, piecewise aggregate approximation*

I. INTRODUCTION

Recently, the field of electrical machines has received increasing attention, towards the development and application of fault detection techniques [1], [2]. Among the most common faults that are encountered in the area of induction motors are: the opening or shorting of one or more of the stator's phase windings [3], the presence of broken rotor bar(s) or cracked rotor's end-rings [4], air-gap irregularities [5], bearing faults, and eccentricity faults [6].

For the detection of these faults and more specifically, for the case of broken bars or eccentricity, various input signals have been used quite successfully. However, methods that rely only on the use of currents, like the (Motor Current Signature Analysis (MCSA) [7], [8] are usually preferred mainly due to their non-invasive nature. The underlying philosophy of those methods is to detect the presence of specific components created by the fault.

For the case of broken bars, the frequencies of the fault component are given by the following equation:

$$f_b = \left(1 \pm 2 \cdot k \cdot s\right) f_s, \qquad k = 1, 2, \ldots \qquad (1)$$

where f_s is the fundamental frequency and s is the slip. Among the various components, the Lower (or Left) Sideband Harmonic (LSH), which is given for $k = 1$ and taking the minus sign in eq. 1: $\left(\left(1 - 2 \cdot s\right) f_s\right)$ [9], is the most extensively used one, in the fault detection literature.

In the case of eccentricity, and more specifically, of mixed eccentricity, which is the most commonly encountered type of eccentricity fault, another group of components, with frequencies given by the following equation, is expected [7]:

$$f_{ecc} = \left(1 \pm m\left(\frac{1-s}{p}\right)\right) f_s, \qquad m = 1, 2, \ldots \qquad (2)$$

As it can be observed, both frequency groups depend on the slip. This observation triggered the development of methods that rely on the analysis of these components during start-up, a period at which the slip varies from one to a value usually close to zero, within a short period of time. This rapid evolution makes the faulty components "draw" quite distinctive patterns in the time-frequency plane. In the case of broken bars, the LSH will show up as a V pattern on the time frequency plane, while in the case of mixed eccentricity and if $m = p/2$, one of the frequency components related to the fault situation, will start at being equal to the supply frequency and will end up in having a frequency just above half the supply frequency (assuming a value of slip close to 0). These characteristic patterns are depicted in Fig. 1.

Various methods under the term Transient MCSA have been proposed for the detection and quantification of the aforementioned components [10]-[14]. Among the different proposed approaches for the diagnosis of these faults or combination of them, the investigation of single faults diagnosis methods has dominated the literature. However, the simultaneous presence of electromechanical faults is the rule and not the exception in industrial settings.

Therefore lately, scenarios where more than one faults can be present at the same time are investigated as in [15]-[17]. In [15], a filter bank combined with high-resolution spectral analysis was used for the estimation of faulty components caused by multiple faults (broken bars, eccentricity, bearing

faults). In [16] a cascade of neural networks was trained to recognize among different fault scenarios, including simultaneous faults. In [17] different faults were investigated through the characteristic "fingerprint" they leave on the decomposed start-up current via Discrete Wavelet Transform (DWT). However this approach has not been yet automated.
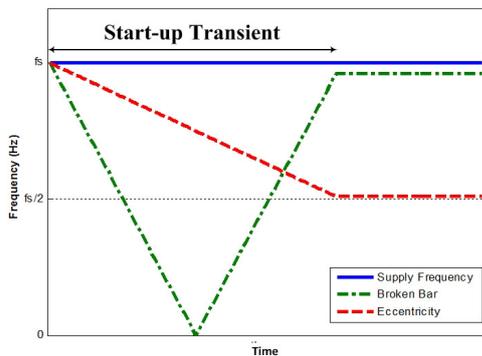

Fig. 1. Approximate evolution of the instantaneous frequency of faulty components during start-up.

In this work the presence of a broken bar as well as mixed-eccentricity faults are examined within the framework of multi-label classification [18], [19]. The proposed novel method uses the time frequency representation of the start-up current as a means for feature extraction. Furthermore, this specific approach, on the one hand has the advantage that allows the aforementioned faulty components to "draw" distinctive pattern on the time frequency plane, while on the other hand poses the challenge of adequately handling the high dimensional representation of the produced time frequency plane. To cope with the latter, a number of steps are involved including the application of a two-dimensional variant of the Piecewise Aggregate Approximation (PAA) and Principal Component Analysis (PCA). The final diagnosis is accomplished within the multi-label classification framework using a very simple binary classifier. The whole procedure is tested using simulations coming from a model of an induction machine, which can accommodate various types of faults.

II. PROPOSED APPROACH

The proposed approach builds upon our previous work on broken bar fault diagnosis using the time frequency representation of the start-up current produced via the short-time Fourier transform (STFT) [20]. That approach is extended to the multi-label case with most of the building blocks remaining the same. For completeness, each stage is presented in brief in the following subsections, while the multi-label stage is presented in more detail since it is new in the field of induction fault diagnosis.

A. Transient Isolation

In this work the isolation of the end of the transient is important. However contrary to the case of broken bars, where at the end of the transient the faulty component gets close and is usually masked by the fundamental/supply frequency, in the

case of mixed eccentricity the faulty component at steady state (assuming a small value of the slip) lies quite apart from the fundamental/supply frequency. Therefore in this work, unlike the case in [20], the window of the current retained for further processing was selected to be approximately 1.5 times the duration of the transient. The end of the transient is detected using a moving window as in [20].

B. Time Frequency Representation

There are many tools that can be used for the analysis of the time-frequency evolution of signals. Among them STFT or window Fourier transform is probably the simplest and most widely used method. Formally is given by Eq. (3), where $w(t)$ represents the window function:

$$X(t,\omega) = \int_{-\infty}^{\infty} x(\tau)w(\tau-t)e^{-j\omega\tau}d\tau \qquad (3)$$

In this work the time-frequency content below the fundamental frequency is considered. For that range, the components created by the broken bar and the eccentricity would have instantaneous frequencies that approximately would look like those in Fig. 1.

As it can be seen in Fig. 2.a which depicts the spectrogram of different fault scenarios (normalized and measured in dB), even though the ideal scenario is not met, both the V pattern and the component created by the eccentricity can be spotted, even though they are a bit distorted.

C. Dimensionality Reduction

The STFT creates a time-frequency representation, practically an image, which has quite high dimensionality. In order to alleviate the "curse of dimensionality" problem two dimensionality reduction techniques are employed.

2D Piecewise Aggregate Approximation

PAA was originally developed for one dimensional signals [21] but since then a two dimensional variant has also been used for reducing the dimensionality of images [20], [22]. In this work 2D PAA is used to transform the high resolution STFT "image" to a more compact representation (Fig. 2.b).

Principal Component Analysis

Even after the application of PAA, the resulting image has a dimension (for this specific implementation) of 10x5, which means that a feature vector of 50 elements will be created. Therefore a second stage is employed to further reduce the dimensionality. In the field of dimensionality reduction many algorithms have been proposed over the past decade. However it turns out that PCA can be quite competitive to other more advanced methods, when it comes to real life problems [23]. PCA transforms the original data by projecting them to just a few of the eigenvectors of their covariance matrix that correspond to the largest eigenvalues [24].

D. Multi-label classification

Standard classification algorithms assume that a given instance (also referred to as example or data point) is

associated with a single label λ_i from a set of disjoint labels Λ [18]. In other words, each instance belongs to just one class λ_i. However in some domains, instances can actually be associated with a set of labels $Y \subseteq \Lambda$, or using a pattern classification terminology, instances can belong to more than one class and thus. these kind of problems fall under the umbrella of multi-label classification [18].
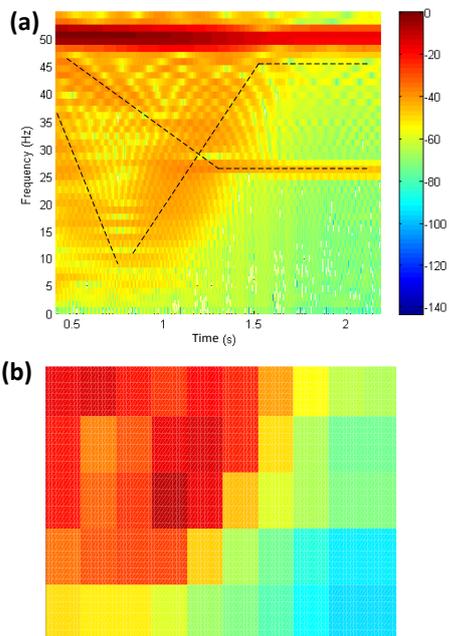

Fig. 2. a)Time-frequency plane of a motor with both mixed-eccentricity and one broken bar, b) 2D PAA

In this case, two basic approaches can be applied: a) problem transformation and b) algorithm adaptation. The former involves the transformation of the multi-label classification problem into one or more single-label classification problems, which can be tackled with any conventional classification algorithm, while the latter, relies on the adaptation of classification algorithms to directly handle multiple labels per instance. In this work the Label Powerset (LP) approach is used [18], which belongs to the first family.

Label Powerset

The LP transformation simply treats each unique *label set* encountered in the training set as one of the classes in a multiclass classification scheme. Therefore, after the transformation, any conventional multiclass algorithm can be used. For example consider the following hypothetical data set $\{\{\lambda_1,\lambda_2\},\{\lambda_1\},\{\lambda_1,\lambda_3\},\{\lambda_3\}\}$. The original dataset has three labels $\{\lambda_1,\lambda_2,\lambda_3\}$ (e.g. in a fault detection framework for induction motors the labels could be {broken bar, eccentricity, bearing faults}). The transformed data will have four classes, and the corresponding set will be $\{\lambda_{1-2},\lambda_1,\lambda_{1-3},\lambda_3\}$. Note that we no longer have sets but single labels – i.e. λ_{1-2} is a single

label that means that this label/class corresponds to labels λ_1 and λ_2 (e.g. {broken bar and eccentricity}) of the original data set.

On one hand, LP has the advantage of taking into consideration potential correlations between the labels. On the other hand it has the disadvantage that it can become computationally intractable for large label sets. For this specific application, due to the small number of classes, computational complexity is not an issue.

E. Minimum Mahalanobis distance classifier

As it was mentioned, the problem transformation methods, have the advantage that common classifiers can be used once the transformed data sets are created. In this work, for the multiclass classification problem that arises in the case of the LP transformation the, Minimum Mahalanobis distance classifiers is used [24], selected for its simplicity and its lack of hyper-parameters that would need extra tuning.

III. METHOD EVALUATION

A. Simulation model

The analytical model used to obtain the simulated waveforms, published in [25], was specifically revised to reproduce the effect of several faults, including electrical asymmetries. Broken bars are simulated by increasing the resistance of this element, up to a level, where the circulating current is negligible, whilst static and dynamic eccentricity are reproduced by modifying the airgap's permeance value. Since saturation is not considered in this model, these values are scaled to approximate the results obtained in the actual equipment.

B. Experimental Procedure - Results

Using the simulation model described in the previous subsection, a motor with two pair of poles is simulated for the following scenarios: a) healthy motor, b) motor with one broken bar, c) motor with mixed eccentricity and d) motor with mixed eccentricity and one broken bar. For each condition ten different start-ups are generated with different durations. For the case of mixed eccentricity, various degrees of static and dynamic eccentricity are simulated within the rage of 5%-35%.

Furthermore, three different values for the number of retained components are tested (two, three and four). The leave-one-out (loo) testing procedure is used and the results in terms of absolute accuracy are summarized in Table I. As it can be seen with the use of only two principal components, not all the relevant information is captured leading to somewhat reduced performance.

TABLE I. CLASSIFICATION RESULTS

PCs	2	3	4
Accuracy	90%	100%	100%

a. PCs-Principal Components

IV. Conclusions

In this work a multi-label classification approach for the diagnosis of multiple faults in induction motors was presented. This is the first time that such an approach is employed for the case of induction motors, especially during the start up. Furthermore, the approach investigated the cases of broken bars and mixed-eccentricity with promising results. Regarding the number of retained principal components, there seems to be a good compromise for this specific setting.

The proposed approach is a data driven approach. Therefore further investigation is needed involving more data as well as data coming from experimental settings. Further investigation is also required regarding the selection of optimal parameters of the feature extraction stages. On the other hand the method is general enough so as to be applied to different scenarios.

In future work the method will be tested using other fault conditions, such as bearing faults and winding shorts, using also experimental data. Moreover different methods from the field of multi-label learning will be examined.

References

[1] A. S. Nandi and H. Toliyat, "Condition monitoring and fault diagnosis of electrical machines-a review," *IEEE Transactions on Energy Conversion*, vol. 20, no. 4, pp. 719–729, Dec. 2005.

[2] G. Georgoulas, M.O. Mustafa, I. P. Tsoumas, J. A. Antonino-Daviu, V. C. Alarcon, C.D. Stylios, and G. Nikolakopoulos, "Principal Component Analysis of the start-up transient and Hidden Markov Modeling for broken rotor bar fault diagnosis in asynchronous machines", *Expert Systems with Applications*, vol. 40, no. 17, pp 7024-7033, 2013.

[3] O. Mustafa, G. Nikolakopoulos and T. Gustafsson, "Stator winding short circuit fault detection based on set membershio identification for three phase induction motors", in *Proc. 20th Mediterranean Conference on Control and Automation*, 2012.

[4] A. P. Santos and T. Lubiny, "A simplified induction machine model to study rotor broken bar effects and fault detection," *European Transactions on Electrical Power*, vol. 20, p. 611, 2010.

[5] A.G.G. Acosta, C.J. Verucchi, and E.R. Gelso, "A current monitoring system for diagnosing electrical failures in induction motors," *Mechanical Systems and Signal Processing*, vol. 20, no. 4, pp. 953–965, 2004.

[6] S. Nandi, T.C. Ilamparithi, L. Sang and H. Doosoo, "Detection of Eccentricity Faults in Induction Machines Based on Nameplace Parameters.", *IEEE Trans. Ind. Electron*, vol. 58, vo. 5, pp 1673-1683, July, 2010.

[7] M. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection", *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 984-993, 2000.

[8] B. M. Ebrahimi, J. Faiz and M. J. Roshtkhari, "Static-, Dynamic-, and Mixed Eccentricity Fault Diagnoses in Permanent-Magnet Syncronous Motors," *IEEE Trans. Ind. Electron*, vol. 56, no. 11, pp. 4727-4739, 2009.

[9] W. T. Thomson, M. Fenger, "Current signature analysis to detect induction motor faults," *IEEE Industry Applications Magazine*, pp. 26-34, July/August 2001.

[10] M. Pineda-Sanchez, M. Riera-Guasp, J. A. Antonino-Daviu, J. Roger-Folch, J. Perez-Cruz, and R. Puche-Panadero, "Instantaneous frequency of the left sideband harmonic during the start-up transient: A new method for diagnosis of broken bars," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4557–4570, Nov. 2009.

[11] J. Pons-Llinares, J.A. Antonino-Daviu, M. Riera-Guasp, M. Pineda-Sanchez, and V. Climente-Alarcon, "Induction Motor Diagnosis Based on Transient Current Analytic Wavelet Transform via Frequency B-Splines," *IEEE Trans. Ind. Electron*, vol. 58, no. 5, pp. 1530-1544, 2011.

[12] G. Georgoulas, M.O. Mustafa, I.P. Tsoumas, J.A. Antonino-Daviu, V. Climente-Alarcon, C.D. Stylios, and G. Nikolakopoulos, "Principal Component Analysis of the start-up transient and Hidden Markov Modeling for broken rotor bar fault diagnosis in asynchronous machines," *Expert Systems with Applications*, vol. 40, no. 17, pp. 7024-7033, 2013.

[13] P. Karvelis, G. Georgoulas, I.P. Tsoumas, J.A. Antonino-Daviu, V. Climente-Alarcon, and C.D.Stylios, "A Symbolic Representation Approach for the Diagnosis of Broken Rotor Bars in Induction Motors," *IEEE Trans. Ind Inf.*, vol. 11, no. 5, pp.1028-1037, 2015.

[14] J. Antonino-Daviu, P. Jover, M. Riera, A. Arkkio, and J. Roger-Folch, "DWT analysis of numerical and experimental data for the diagnosis of dynamic eccentricities in induction motors," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp.2575-2589, 2007.

[15] A. Garcia-Perez, R. de Jesus Romero-Troncoso, E. Cabal-Yepez, R.A. Osornio-Rios, "The Application of High-Resolution Spectral Analysis for Identifying Multiple Combined Faults in Induction Motors", *IEEE Trans. Ind. Electron*, Vol. 58, No. 5, 2011.

[16] V.N. Ghate, and S.V. Dudul, "Cascade neural-network-based fault classifier for three-phase induction motor," *IEEE Trans. Ind. Electron*, vol. 58, no. 5, pp.1555-1563, 2011.

[17] J. Antonino-Daviu, P. Jover Rodriguez, M. Riera-Guasp, M. Pineda-Sanchez, and A. Arkkio, "Detection of combined faults in induction machines with stator parallel branches through the DWT of the startup current," *Mechanical Systems and Signal Processing*, vol. 23, no. 7, pp. 2336-2351, 2009

[18] G. Tsoumakas, I. Katakis, and I. Vlahavas, I., 2009. Mining multi-label data. *In Data mining and knowledge discovery handbook*, pp. 667-685. Springer US, 2009.

[19] C.M. Vong, P.K. Wong,. And W.F. Ip, "A new framework of simultaneous-fault diagnosis using pairwise probabilistic multi-label classification for time-dependent patterns," *IEEE Trans. Ind. Electron*, vol. 60, no. 8, pp.3372-3385, 2013.

[20] G. Georgoulas, P. Karvelis, C.D. Stylios, I.P. Tsoumas, J.A. Antonino-Daviu, and V. Climente-Alarcon, "Automatizing the broken bar detection process via short time Fourier transform and two-dimensional piecewise aggregate approximation representation," *In Energy Conversion Congress and Exposition (ECCE)*, pp. 3104-3110, 2014.

[21] B. K ,Yi and C. Faloutsos, "Fast time sequence indexing for arbitrary Lp norms," *in proc. of the VLDB*, 2000.

[22] J. Chen, Y. S. Moon, M. F. Wong, and G. Su, "Palmprint authentication using a symbolic representation of images," *Image and Vision Computing*, vol. 28, no. 3, pp. 343-351, 2010

[23] L.J. van der Maaten, E.O. Postma, H.J. van den Herik, "Dimensionality reduction: A comparative review," *Journal of Machine Learning Research,* vol. 10 pp. 66-71, 2009.

[24] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2002.

[25] M. Pineda-Sanchez, V. Climente-Alarcon, M. Riera-Guasp, R. Puche-Panadero, "Enhanced Simulink Induction Motor Model for Education and Maintenance Training," *Int. Journal of Systemics, Cybernetics and Informatics*, vol. 10, no. 2, 2012