Expert Systems with Applications 41 (2014) 6386–6399

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Efficient residuals pre-processing for diagnosing multi-class faults in a doubly fed induction generator, under missing data scenarios

Roozbeh Razavi-Far [a,*], Enrico Zio [a,b], Vasile Palade [c]

[a] Department of Energy, Politecnico di Milano, via Ponzio 34/3, 20133 Milan, Italy
[b] Systems Science and the Energetic Challenge, European Foundation for New Energy-Electricité de France, Ecole Centrale Paris and Supelec, Paris, 92295 Chatenay-Malabry Cedex, France
[c] Faculty of Engineering and Computing, Coventry University, Priory Street, Coventry, UK

ARTICLE INFO

Keywords:
Fault diagnosis
NIPALS
Wold cross-validation
Latent residuals
New class faults
Wind turbine

ABSTRACT

This paper focuses on the development of a pre-processing module to generate the latent residuals for sensor fault diagnosis in a doubly fed induction generator of a wind turbine. The pre-processing module bridges a gap between the residual generation and decision modules. The inputs of the pre-processing module are batches of residuals generated by a combined set of observers that are robust to operating point changes. The outputs of the pre-processing module are the latent residuals which are progressively fed into the decision module, a dynamic weighting ensemble of fault classifiers that incrementally learns the residuals-faults relationships and dynamically classifies the faults including multiple new classes.

The pre-processing module consists of the Wold cross-validation algorithm along with the non-linear iterative partial least squares (NIPALS) that projects the residual to the new feature space, extracts the latent information among the residuals and estimates the optimal number of principal components to form the latent residuals. Simulation results confirm the effectiveness of this approach, even in the incomplete scenarios, i.e., the missing data in the batches of generated residuals due to sensor failures.

1. Introduction

The doubly fed induction generator (DFIG) is one of the most widely used classes of induction machines in the megawatt-class wind turbines (Hansen & Michalke, 2007). The DFIGs have shown a good performance in normal operation, but they are quite sensitive to particular classes of faults. The rate of failures in the sensors and the generator of wind turbines are reported to be approximately 14.1% and 5.5% of the total number of failures, that cause 5.4% and 8.9% of the system downtime (Ribrant & Bertling, 2007).

The sensor fault detection and isolation in the DFIGs has an important role to guarantee the safe and reliable operation of wind turbines. Since monitoring the generator entails processing the current and voltage sensor measurements, the first step is devoted to sensor fault diagnosis, which has been addressed in recent works (Boulkroune, Galvez-Carrillo, & Kinnaert, 2010; Galvez-Carrillo & Kinnaert, 2010).

Fault diagnosis can be performed in two major steps. Firstly, several signals, so-called residuals, reflecting faults in the process behavior, are generated. In the second step, the residuals are evaluated for decision making, to determine the time and the location of potential faults (Razavi-Far, Davilu, Palade, & Lucas, 2009a, 2009b).

Multiple observers schemes were developed in Boulkroune et al. (2010) and Galvez-Carrillo and Kinnaert (2010) to generate residuals associated to stator voltage and current sensors, as well as rotor sensors, respectively. These multiple observers were integrated in Razavi-Far and Kinnaert (2012, 2013) to reveal the mutual effects of the faults in each type of sensors on the residuals associated to another sensor type. This coupling prevented to develop a decision system by basic combination of the previously developed decision systems for each class of sensors. Thus, an effective classification technique has been used to design a suitable decision system in Razavi-Far and Kinnaert (2012, 2013).

The problem of fault classification can be tackled resorting to computational intelligence techniques. However, these approaches are usually based on time-series data of various signals in static environments (Razavi-Far, Davilu, Palade, & Lucas, 2009b). On the contrary, in dynamic environments, an incremental learning

* Corresponding author.
E-mail addresses: roozbeh.razavi@gmail.com, roozbeh.razavi@mail.polimi.it (R. Razavi-Far), enrico.zio@polimi.it, enrico.zio@ecp.fr, enrico.zio@supelec.fr (E. Zio), vasile.palade@coventry.ac.uk (V. Palade).

http://dx.doi.org/10.1016/j.eswa.2014.03.056
0957-4174/

strategy is needed to update the decision system for fault classification. This has been done by resorting to ensemble of fault classifiers (Baraldi, Razavi-Far, & Zio, 2011b). In Baraldi, Razavi-Far, and Zio (2011a), a bagged ensemble of Fuzzy C-Means (FCM) classifiers was used for fault classification and its confidence for decision making has been studied in Baraldi, Razavi-Far, and Zio (2010).

The incrementally trained ensemble of classifiers in Baraldi et al. (2011b) can learn the new relations between the upcoming signals, while keeping the previously trained classifiers to preserve the existing knowledge. Albeit this has been successfully applied for decision making and fault classification in changing operating conditions (Baraldi et al., 2011b), the situation becomes more complicated when the datasets collected in subsequent installments have patterns of new classes of faults that were not included in previous datasets. Consequently, the base classifiers of the ensemble are doomed to misclassify patterns from faulty classes on which they were not trained.

The problem of new class fault diagnosis was firstly tackled by resorting to dynamic weighting ensembles (Razavi-Far, Baraldi, & Zio, 2012a), where a dynamic weighting ensembles algorithm was adopted for fault diagnosis in the feedwater system of a boiling water reactor (BWR). The algorithm is particularly developed for incremental learning of multiple new concept classes of faults. The detection of unseen classes in subsequent data was based on thresholding the normalized weighted average of the outputs (NWAO) of the base classifiers in the ensemble (Razavi-Far et al., 2012a, Razavi-Far, Baraldi, & Zio, 2012b).

Here a multiple observer scheme is used for residual generation, while for residual evaluation a dynamic weighting ensemble of classifiers is used. In the first step, a bank of observers generates a set of residuals that are robust to operating point changes. A so-called signal-based approach is used for residual generation of the stator current and voltage sensors, while two-stage filters exploiting the DFIG model and the balanced signal model are used for residual generation of the rotor currents, the same as in Razavi-Far and Kinnaert (2012, 2013).

In the second step, the pre-processed residuals are progressively fed into the dynamic weighting ensembles for fault classification. The algorithm incrementally learns the relation between projected residuals and faults, and dynamically classifies the faults including multiple new classes.

In Razavi-Far and Kinnaert (2012, 2013), prior to fault classification, the generated residuals ($r_i = r_1, r_2, \ldots, r_9$) were resampled (i.e., down-sampled) in the processing module and then forwarded to the fault classifier, i.e., the 'second step'. Each residual contains two vectors that form 18 features for the dynamic weighting ensemble of fault classifiers. The dynamic weighting ensemble of fault classifiers mapped these 18 features to 10 possible classes. These classes include the normal state '*ff* or fault-free' and 9 classes of faults (f_i=$f_1, f_2, \ldots, f_9$). The first three faults are sensor faults in the stator voltage at phase (a, b, c). Other faults correspond to sensor faults in stator and rotor currents at phase (a, b, c), respectively. In the preceding works (Razavi-Far & Kinnaert, 2012, 2013), it was shown that the decision module of the diagnostic system can isolate all classes with respect to the unavailability of patterns from all faulty classes during the training (i.e., new faults became available dynamically in the course of time). The major focus was on detection and isolation of additive step-like faults, but additive drift-like faults were taken into account as well.

The generated residuals by multiple observers contain redundant or irrelevant residual vectors (i.e., features) that can degrade the fault classification performance. The pre-processing module in Razavi-Far and Kinnaert (2012, 2013) only resamples (i.e., down-samples) the residual vectors, which is a pattern-wise process.

To improve the fault classification performance, a pre-processing of the features is necessary. Feature selection is a task of pre-processing the data to select a subset of features. Feature extraction generates new features (e.g., latent residuals) from functions of the original features (i.e., generated residuals by multiple observers). This can improve the fault classification performance by improving the model interpretability, reducing overtraining, enhancing the generalization capability and shortening the training times.

The feature selection methods can be divided into three main categories: wrappers, filters and embedded methods (Guyon & Elisseeff, 2003). There exist different methods for feature extraction, such as those presented in Vong and Wong (2011), Vong, Wong, and Ip (2013) and Bruzzese (2014).

Moreover, the fault classifiers of the diagnostic system, like any other type of classifiers, fail to classify the incomplete patterns (i.e., containing some missing features). Thus, it is necessary to discard or impute the patterns with missing data/features before sending to the dynamic weighting ensemble. In the first case, the fault classification module cannot classify the fault for the missing patterns and the final decision is in question. In the latter case, the missing data are imputed in advance and, thus, the missing patterns can also be classified. There exists different number of missing data imputation techniques (Gheyas & Smith, 2010; Rassler, Rubin, & Zell, 2013). Although, these methods can impute the missing data, the outcome is the completed dataset that needs further pre-processing to reduce the size of features. This can be computationally expensive, and not feasible for online monitoring and diagnostic applications. Therefore, here a non-linear iterative partial least squares (NIPALS) algorithm along with the Wold cross-validation (Wold, 1978) has been used for pre-processing. This algorithm is fast and suitable for online application, it extracts the latent variables (i.e., latent residuals) from the residual datasets, and it handles the missing data.

This paper aims to study the residuals and focus on the pre-processing module to provide more informative features of smaller size for fault classification. An efficient way to process the generated residuals is developed in order to extract latent information among residuals and provide informative features for the decision module of the diagnostic system. This is done by resorting to the principal component analysis (PCA), a popular data analysis technique.

The contribution of this work is in developing a non-linear iterative partial least squares (NIPALS) algorithm along with a dynamic weighting ensemble, for residual evaluation and new class fault diagnosis in dynamic environments. This algorithm is capable of reducing the number of the generated residuals, which incrementally become available, and projecting them onto the new feature space of smaller size, by extracting the latent information. The projected latent residuals allow faster incremental update of the ensemble of fault classifiers and improve the classification accuracy of some of the faults, while incomplete batches of residuals become available. The proposed classification scheme is validated on the problem of early diagnosis of new class faults in the sensors of a DFIG.

The rest of this paper is organized as follows. Section 2 describes briefly the system and presents the fault diagnostic scheme with a focus on the pre-processing module. Section 3 presents the NIPALS algorithm for the dimensional reduction of incomplete data. Next, the Wold cross-validation algorithm is used along with the NIPALS algorithm to estimate the number of principal components and extract the latent residuals. In Section 4, an application to the sensors of DFIG-based wind turbines is presented. First, the generated batch of residuals by multiple-observers are processed by the Wold cross-validation along with the NIPALS algorithm to form the latent