Electric Power Systems Research 166 (2019) 1–8

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Performance analysis of a learning structured fault locator for distribution systems in the case of polluted inputs

E. Correa-Tapasco, J. Mora-Flórez*, S. Perez-Londoño

Electrical Engineering Program, Universidad Tecnológica de Pereira, Pereira, Colombia

ARTICLE INFO

Keywords:
Fault location
Learning based fault locators
Distorted signals
Performance

ABSTRACT

Waveform distortion and the service continuity are the two main aspects which define the power quality; in the last, the fault location plays a fundamental role. Considering the requirement of reliable tools to improve the distribution system operation, this document analyses the confidence of a fault locator approach. This paper initially presents the adjusting and the validation methodology used to develop a fault locator, and also its performance is analysed by considering different distorted measurements as inputs. The approach here presented is a learning based fault locator (LBFL), specially customized for distribution systems, whose core are the Support Vector Machines (SVM). The LBFL is tested in an unbalanced 34.5 kV distribution feeder considering diverse operating conditions and fault scenarios, which include several fault resistances, locations, and common disturbances on the voltage and current input signals. From the tests, the lowest LBFL performance is 96.5% in such cases of not distorted measurements of voltage and current. In the case of considering polluted inputs, the worse performance is obtained in the case of distortions caused by low sampling frequency, current transformer saturation and high harmonic distortion.

1. Introduction

Service discontinuity is a common problem in electricity distribution networks, where system design and the fault management are in the core as solution alternatives [1–3]. As a consequence, relevant research efforts have been oriented to the development of fault location methods, which are proposed to assist utility operators in expediting service restoration, and consequently reducing outage time and cost [4–9]. The survey presented in Ref. [3], classifies the fault location methods in these based on the estimation of the fault impedance [10], strategies which use measurements taken along the system and the classical circuit analysis [11], approaches based on traveling waves [12], methods which use the learning theories [4–9,13] and approaches based on the combination of the previous strategies [14,15].

The learning-based fault locators (LBFL), as the here analysed, use databases composed by the fault location and the voltage and current phasors, obtained during prefault and fault steady stages, and are aimed to establish a relation between this information. These learning based relation, which in this case is called support vectors, is next used to obtain the fault location in the case of analysing new faults. These methods determine the faulted zone and contribute on the elimination of multiple estimation of the fault location, which is a common problem in fault location approaches [10].

In most of the analysed references, the effects of distorted fault databases on the learning bases fault locator performance are not considered and then the confidence in such cases remains unknown [5,15–18]. This is an important aspect to be strictly defined, considering that in most of the cases, the measurements of voltages and currents used to define the databases are distorted.

According to the exposed, this paper is focused on the implementation and evaluation of a LBFL, considering several waveform distortion cases. The analysed distortions are originated by the measuring devices or by the electromagnetic phenomena [1–2,19–27]. Considering the above exposed, this research is oriented to define the minimal requirements of a LBFL, on the basis of the desired performance and in the case of fault location applications in distribution systems.

As contents, this paper is divided in four sections: the fundamental theoretical concepts are briefly described in Section 2. Then, the LBFL validation strategy used to adjust and to determine the performance in several distortion cases is proposed at Section 3. Test and discussion are presented in Section 4. Finally, the main conclusions are presented in the last section.

* Corresponding author.
E-mail addresses: ever@utp.edu.co (E. Correa-Tapasco), jjmora@utp.edu.co (J. Mora-Flórez), saperez@utp.edu.co (S. Perez-Londoño).

https://doi.org/10.1016/j.epsr.2018.09.016
Received 2 February 2018; Received in revised form 10 August 2018; Accepted 18 September 2018

2. Conceptual background

This section is devoted to analyse the main concepts required in this research. Complete explanations relate to these concepts are at the references.

2.1. Support Vector Machines as the core of the LBFL

According to several references as Refs. [4–8], one of the best performance learning based algorithms to data analysis is the known as Support Vector Machines (SVM), which is based on quadratic programming, several constrains and kernel transformations. This algorithm has been successfully used in many classification and regression problems, therefore the use of SVM as classification technique to assist fault location is here considered [5–9,13]. The main characteristic of the SVM are next described.

2.1.1. SVM in case of linear cases

The SVM basis comes from the statistical learning theory and these are defined initially as a binary classification technique. SVM are also derived from the optimization of the separation threshold defined by the perceptron in neural networks, the use of kernel methods and the generalization theories [28]. Classification using SVM requires of training and testing data sets. In the training set, each data consists of x_i attributes in a N dimensional space and a target value called class (usually 1 or -1 which in this case corresponds to two faulted zones), as in Eq. (1) [7–9].

$$x_i \in R^N \text{ and } y_i \in \{+1, -1\} \tag{1}$$

The classifier helps to develop a model, which can predict the class label of unknown data (faulted zone), considering a confidence indicator of performance. Although the SVM algorithm is defined as a binary classification, this is used to classify faults in more than two classes using decomposition and reconstruction methods [28,29]. The SVM basic algorithm determines an optimal separation hyperplane H given by $y = \boldsymbol{w} \cdot x + \boldsymbol{b} = 0$, which has the maximum margin to the training nearest pattern, forcing the generalization of the learning machine as shown in Fig. 1 [6,28]. Weight ($\boldsymbol{w}$) and bias ($\boldsymbol{b}$) control the function and those data points that the margin pushes up against are called "support vectors" (k, l, m and n). To find the optimal separation hyperplane, the optimization problem presented in Eq. (2) is solved, considering that margin is inversely proportional to $(w{\cdot}w)^{1/2}$ [9,29].

$$\min_{w,b}\frac{1}{2}(w{\cdot}w), \quad \text{subject to}: y_i(w{\cdot}x_i + b) \geq 1, \forall\, i \tag{2}$$

2.1.2. SVM considering soft margin

The basic SVM formulation considers the nonexistence of mixed classes (zones). To deal with this, the SVM are reformulated by considering relaxation variables (ξ) to define the "soft margin". Thus, the optimization problem presented in Eq. (2) is now given as Eq. (3), where C is the error penalization constant [29].

$$\min_{w,b}\frac{1}{2}(w{\cdot}w) + C\sum_{i=1}^{n} \xi_i$$
$$\text{subject to}: y_i(w{\cdot}x_i + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0, \forall\, i \tag{3}$$

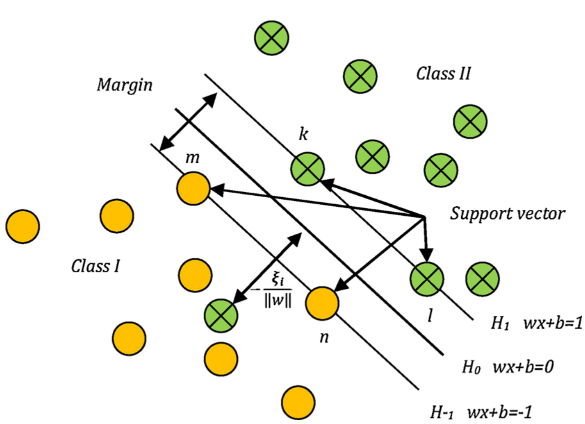

Fig. 1. Separating hyperplanes in a SVM.

2.1.3. SVM in non-linear cases

To consider non-linear separable classes (zones), it is possible to transform the input into a new higher dimension space, where these are linearly separable. The transformation function $F(\cdot)$ is defined as the inner product of the input data in the original classification space. Linear SVMs are extended to non-linear cases by using an appropriate kernel function [28–30]. When a Radial Basis Function (RBF) is chosen as kernel function, two parameters (constant C and kernel parameter $\boldsymbol{\sigma}$) are required. In this paper, the Gaussian RBF kernel presented in Eq. (4) is used. At this paper, the cross validation is used to select the best combination of C and $\boldsymbol{\sigma}$ [28].

$$F(x, y) = exp\left(-\frac{\|x - y\|^2}{2\boldsymbol{\sigma}^2}\right) \tag{4}$$

2.2. Power quality disturbances

The considered power quality disturbances are the alterations on the RMS magnitude, transients and distortions at the waveform [27]. The considered disturbances are originated by low sampling frequency (*sf*), saturation of measuring transformers, resolution of the measuring devise (*R*), harmonics (*H*), noise (*N*) and the presence of DC offset component (*DCO*) [2,17–27].

The effect on the LBFL performance of the waveform distortion sources is here analysed as the main part of the proposed research.

3. LBFL validation strategy

The LBFL uses the SVM as a learning tool to determine the fault location as the result of a classification task. The validation methodology proposed to analyse the performance of the LBFL is divided in three stages, as is depicted in Fig. 2. The first one is oriented to determine the descriptor database from the measurements of current and voltage obtained during fault situation; the second stage is oriented to adjust the LBFL by the iterative application of training and testing processes. Finally, the third stage is devoted to the performance comparison of the LBFL by considering inputs composed by distorted and non-distorted measurements of voltage and current.

3.1. Stage 1. Determination of the descriptor database

Three steps compose this stage, as is shown in Fig. 3. The first one is the zone definition at the analysed distribution system; the second is the conformation of the fault database composed by the measurements of current and voltage during faults and considering several operating conditions of the distribution system. The last step is oriented to obtain the descriptor database from the fault database. A descriptor is here

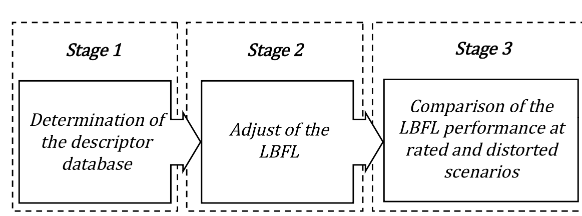

Fig. 2. Proposed fault locator validation methodology.