Discriminating Grounding Systems Configurations using the Pattern Recognition Framework

Francisco Alexandre A. Souza, *Student Member, IEEE,* and Tobias R. Fernandes Neto and Rodrigo P. Magalhães and Felipe Bandeira Silva and Rooney R.A. Coelho and Ricardo Silva Thé Pontes

***Abstract*—This paper presents an innovative method for inspecting grounding systems based on the impulse response of the grounding system and the pattern recognition framework. Such kind of automatic tool can be of valuable help while inspecting grounding systems, since it allows the inspection of the grounding systems already covered or cemented, to check whether they are arranged properly or not. Moreover, the proposed system is easy to mount and take few seconds to evaluate the grounding configuration. The proposed system is composed by four main elements, an excitation system, a data acquisition, a feature extraction and a pattern recognition model. To evaluate the proposed system, four grounding configurations are considered to be detected in the experimental part. The results demonstrate that the proposed system can distinguish with up 70% among these four configurations.**

I. Introduction

Grounding systems (GS) play a vital role in electrical systems. The correct design of GS are of vital importance for the safety and protection of personnel, equipments and facilities, as well to the correct operation of the electricity supply network, etc. The selection of correct configuration, i.e. number and disposition of rods, for the GS is dependent on several factors, which includes the soil resistivity, the available area to install the GS, the season-weather characteristics of the region, the demands of the project, etc [1]. Once the configuration of the GS to be installed is defined, its installation should be properly implemented. However, in some cases, this does not happen, mainly because of errors during the project execution, such as installing an incorrect number of rods (usually less than specified) and/or by forgetting to properly connecting the rods. This situation has a small probability to happen, but for a large amount of installations its effect can be very significant, since in some situations it can be the cause of interruptions in electrical distribution, damage electrical equipments, etc.

The problem of incorrectly installing the GS can be minimized by visually inspecting the GS after its installation. However, there is the possibility that the GS have already been covered or cemented, thereby hindering a visual inspection of the installed system. Under this kind of situation it would be beneficial to have a method, independent of visual inspection and easy to mount, to check whether the configuration of GS is arranged properly or not. To tackle this problem, a new methodology for inspecting GS was developed in this paper. It is based on the analysis of the impulse response of the GS. Moreover, it is a portable device. The proposed method is able to identify the number of ground rods in the GS when it is installed.

Several works have attempted to study the behavior of the GS when submitted to a lightning strike, [2]–[8] these approaches have modeled the GS response based on circuit theory, transmission line theory, full wave models based on field equations and the Numerical Electromagnetic Code. From the literature [2]–[8] it is possible to conclude that there are two main elements that affect the GS response: the soil and ground rod properties. Clearly, different GS configurations produces different responses. However, modeling different GS configurations based on first-principle models can be difficult and so far it has not been investigated in literature. The main objective of this work is to distinguish between different GS configurations, and for that purpose a data-driven model, based on pattern recognition framework is employed. The use of data-driven approach overcome the complexity and the time demanding of first-principle models. In the proposed approach, the pattern recognition model is trained using the GS response of different configurations from different soils. It has shown to be able to distinguish between different GS configurations. From the machine learning perspective, the objective of checking whether the configuration of grounding is arranged properly or not can be viewed as multiclass or binary classification problem, where in the later, the correct configuration is considered as the positive class, and the incorrect configuration is considered as the negative class. For example, assume that the configuration to be mounted should be three parallel rods, then possible incorrect configurations are two rods, one rods and no rod. A number of rods greater than the specified rarely happens, then this cases will not be considered here.

The proposed approach is subdivided in four subsystems: excitation, acquisition, feature extraction and pattern recognition subsystems. The excitation subsystem is responsible to inject a impulse voltage signal similar to a lightning stroke into the installed grounding system. The acquisition subsystem is responsible to acquire the response signals (voltage and current) at a high sample rate rate of 2MSa/s; this high sample rate is required to capture the transient response of the system. The feature extraction module is responsible to extract the features from the transient signal of the impedance, determined through the voltage and current responses. The extracted features are the fast Fourier transform (FFT) coefficients (harmonics) of the

Francisco Alexandre A. Souza, and Tobias R. Fernandes Neto, and Rodrigo P. Magalhães, and Felipe Bandeira Silva, and Rooney R.A. Coelho and Ricardo Silva Thé Pontes; Department of Electrical Engineering, Laboratory of Energy Efficiency in Motor Driven Systems - LAMOTRIZ, Federal University of Ceará, Fortaleza - CE, Brazil

impedance signal. The extracted features are used as input of a pattern recognition model in the classification subsystem, that outputs the information regarding whether the configuration of grounding is arranged properly or not.

In order to evaluate the proposed grounding inspection system, a set of 26 controlled experiments, in distinct soils were conducted. The data collection was done in different seasons and different locations. In these experiments, the configurations with one, two, three and four ground rods in horizontal line were mounted and the impulse response was acquired. All the ground rods have $2.4m$ length with $0.0150m$ diameter, and for the configurations with two, three and four rods, the rods were spaced in intervals of $3m$. Several pattern recognition models were evaluated according with its capability of detecting different GS configurations. The pattern recognition models tested were the Adaptive Boosting (Adaboost), support vector machine (SVM) and random forest classifiers (RF). The proposed approach has show to be able to inspecting grounding system and determine the configuration of grounding rods with a moderate rate of accuracy, up to 70% for some GS configurations.

This paper is organized as follows. Section II gives the motivation of the proposed work. Section III describes how the problem of determining different grounding configurations is inserted into the pattern recognition framework. Section IV describes the proposed methodology for inspecting GS. Section V presents the experimental results. Finally, Section VI gives the conclusion remarks.

II. Motivation

In some Brazilian utilities a single rod ($2.4m$ long and with $0.0150m$ diameter) is frequently applied throughout medium-voltage lines at the service entrances of low-voltage consumers, and to provide periodical grounding points. In other cases, as in pole-mounted distribution transformers protected by surge arresters, the typical grounding configurations applied are three parallel rods in straight line ($2.4m$ long with $0.0150m$ diameter), spaced in intervals of $3m$ [9]. In most cases, the installation of such medium voltage lines and the corresponding GS is made by third party companies and the verification of GS is done visually, which in most cases is inefficient due to the fact that the GS is covered or cemented. In this context, there are reports showing that there are installations with an incorrect number of rods (typically less than specified). Moreover, the ground rods are likely to degrade over the time due to exposure to the earth environment causing its corrosion, and actually it is not possible to verify its consistency. The proposed method can also be used as a method to help the diagnose of ground rods.

Having a portable device, easy to mount and with fast capability of inspecting the GS is of great importance for the maintenance and inspection of GS. In this work, such methodology is described and it has been evaluated with respect its ability to detect one, two, three and four ground rods in horizontal line.

III. The Pattern Recognition Framework and the Discrimination of the Grounding Systems Configurations

The objective of the pattern recognition framework is to find a mapping function (classification model):

$$h(\mathbf{x}, \boldsymbol{\zeta}) : X \rightarrow Y, \tag{1}$$

that maps the input domain X into the output domain Y. An element of X is represented by the vector $\mathbf{x} \in \Re^n$, and the an element of Y is given by the classes $y = \{y_1, y_2, \ldots, y_c\} \in \Re^c$, where c represents the number of classes. If $c = 2$ it is considered as a binary problem and if $c > 2$ it is considered as a multiclass problem. In the binary classification problem one class is considered as the positive class, and the other class as the negative. The major issue of pattern recognition framework is to select the parameters $\boldsymbol{\zeta}$ of the classification model appropriately. This is done by training the model h with a set of N examples $\Phi = \{(\mathbf{x}_i, y_i) | i = 1, \ldots, N\}$. The selection of the best value of $\boldsymbol{\zeta}$ is done in the model learning phase.

In fact, the problem of inspecting a GS can be done by using the pattern recognition framework in two ways, by means of multiclass and binary classification models. These approaches are described as follows.

Under the assumption that $C > 2$, i.e. with the objective of discriminating more than two GS configurations, the inspection of any grounding system can be naturally viewed as a multiclass classification problem, where the input $\mathbf{x}$ in (1) is a extracted characteristic of the GS to be determined and the output y_i for any $i \in 1, \ldots, c$ represents the GS configuration. The values of y_c for $c = 1, \ldots, C$, represent different GS configurations.

In the multiclass classification, it is necessary to train only one model, since it gives the GS configuration as output. The pattern recognition framework can be a excellent tool to help evaluating GS in a fast and reliable way. The major drawback when compared with the physical modeling is the necessity to have a set of examples to learn the model. These examples should be collected in different soils, as many as possible, and the GS configurations which are going to be inspected/compared should be mounted and their responses should be acquired and stored. These response can be acquired by using a excitation system and a set of measurement parameters. Afterwards, by using the stored experiments it is possible to learn a set the pattern recognition models and find the best one.

IV. Proposed Methodology

To compose the proposed system two steps are obligatory: data collection and training data set definition, and the model definition. The system can evaluate any GS configuration when the obligatory steps are done. The proposed system is shown in Fig. 1. All these steps share four main systems: an excitation, a data acquisition, a feature extraction and a pattern classification and model learning systems. As the obligatory steps make use of the four main systems, they are described first than the obligatory steps as follows.

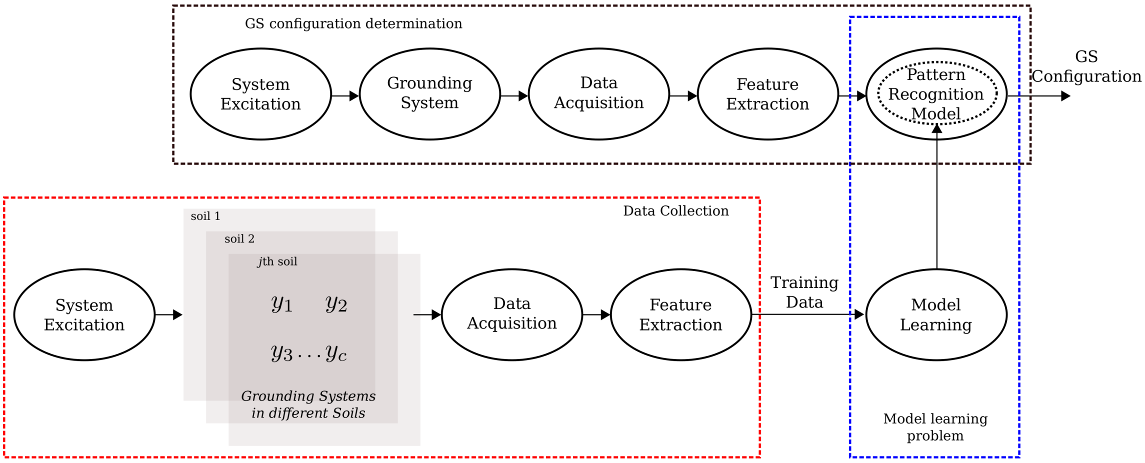


Fig. 1. Architecture of the proposed grounding system discrimination methodology.

A. Shared Systems

In this section the four main systems are going to be discussed. In each of the following sections the purpose of the system and the solution used in this paper will be given.

1) Excitation System: The objective of excitation system is to enable the acquisition of the transitory of the GS response, which is going to be used as basis to discriminate the different GS configurations; the transitory contains the RLC characteristic of the ground-rod set. The transient response of the GS can be extracted by injecting a voltage signal similar to a lighting stroke and collecting the impulse response (the data collection will be discussed in Section IV-A2). According to [10, Chapter 6] the lighting stroke waveform can be approximated by a double exponential as follows:

$$v(t) = V_0 \left(e^{-\alpha t} - e^{-\beta t}\right), \tag{2}$$

where V_0 is the peak value and α and β are constants that define the front time t_f and tail time t_t, respectively. According to [10], the following approximations are used in literature:

$$t_f \approx \frac{1}{\beta}, \tag{3}$$

$$t_t \approx \frac{1}{\alpha}. \tag{4}$$

The common values of the front time of a lighting strikes is between $0.5\mu s$ to $10\mu s$, while its peak value decays to 50% after $30\mu s$ to $200\mu s$. Fig. 2 shows the lighting strike waveform for different values of α and β and $V_0 = 1000$V, the values of α and β range in the common values of a lighting stroke. The excitation system circuit, proposed in this paper, is depicted in Fig. 3. This circuit can approximate the typical waveform of lighting stroke, given by (2). The system of Fig. 3 can be described by two main steps. The first one is the capacitor charging C_{out} step and the second one is the application of the capacitor voltage into the GS. In the first step, the MOSTFET G_1 is closed and G_2 is kept open. The transformer TR step-up the voltage from $220V$ to $1kV$, approximately. Then, the voltage is rectified, by the full bridge rectifier composed by the diodes D_1, D_2, D_3, D_4 and capacitor C_{link}, so that the output capacitor C_{out} is charged, the role of resistor R_1 is

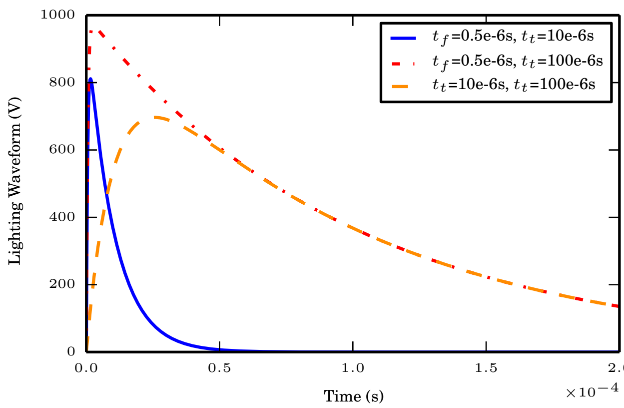


Fig. 2. Lighting stroke waveform for different value of β and α.

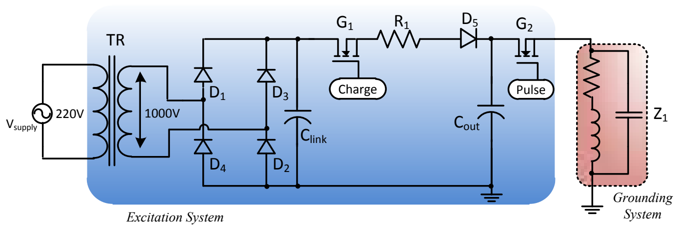


Fig. 3. Architecture of the proposed excitation system.

to limit the current to the capacitor C_{out}. In the second step, the MOSFET G_1 is opened and MOSFET G_2 is closed, then the output capacitor C_{out} (charged before in the first step), will generate an impulse voltage over the GS impedance Z_1 through the MOSFET G_2 that is closed.

It was demonstrated experimentally that the behavior of the excitation system in Fig. 3, has a similar behavior of a double exponential, given by (2). The voltage applied to the grounding system is in the order of 1kV. As the circuit of Fig. 3 is uncontrolled, the parameters t_f and t_t of the double exponential waveform are dependent of the soil properties. Fig 4, show experimental results of the voltage waveform applied in different soils.

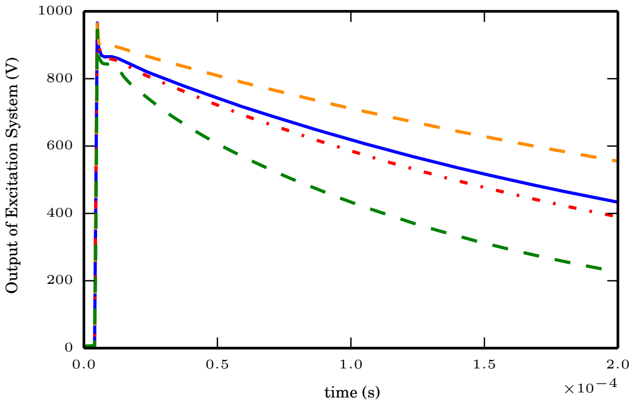


Fig. 4. Output of the Excitation System for different soils.

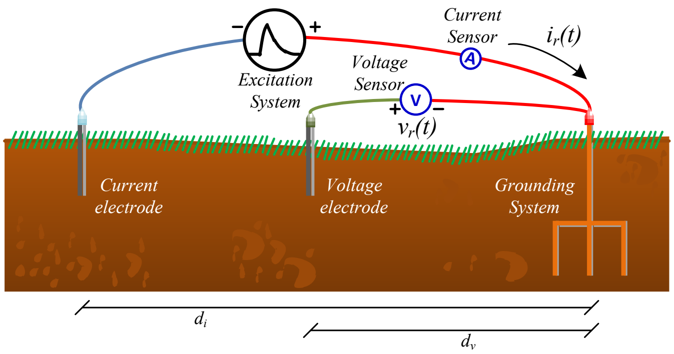


Fig. 5. Architecture of the proposed excitation system.

2) Data Acquisition System: The data acquisition is responsible to define and acquire the response data after applying the excitation signal into the soil.

In this paper, two signals are acquired as responses. The first one is the voltage signal, $v_r(t)$, and second one is the current signal, $i_r(t)$, responses. These signals are collected by two auxiliary electrodes, in line, placed from distances d_v and d_i from the GS under test, Fig. 5 shows the structure of the proposed system. This is based on the 3-point method fall of potential test for measuring the resistivity of the soil, but differently of this methodology the value of d_v is kept fixed in all experiments. In this paper, the distances d_v and d_i were defined as $d_v = 12.5$m and $d_i = 20$m. The current and voltage electrodes were disposed in straight line in all experiments. These distances were defined empirically. However, others can use different values for these distances, under the constraint that it should keep it fixed for all experiments.

The transient response, which is assumed to contain the information necessary to discriminate different grounding configurations, is in the order of micro seconds (μs). To allow the acquisition of the transient signal, this work acquires the data at a sample rate of 2MSa/s, by using the data acquisition system U2531A from Agilent. This sample rate is enough to capture data samples in the order of micro seconds (μs). The voltage and current values were acquired using two Hall Effect transducers from LEM.

3) Feature Extraction System: After the voltage $v_r(t)$ and current $i_r(t)$ responses are acquired, by the data acquisition system, the transient of the impedance in time domain, given by

$$z_r(t) = \frac{v_r(t)}{i_r(t)}, \tag{5}$$

is determined. The transient response corresponds to the first $125\mu s$, which is equivalent to the first 250 samples, under the rate of data acquisition of 2MSa/s. To extract relevant information of the response signal, the Fast Fourier transform (FFT) is applied on $z_r(t)$ for $t = 1, \ldots, 250$. Due to the characteristic of FFT symmetry, only the first 125 coefficients (harmonics) are considered as feature for input in the classification model. Then, the input sample $\mathbf{x}_i$, for the GS represented by y_c are the first 125 FFT coefficients of its response. In the case of this paper $\mathbf{x}_i \in \Re^{125}$.

The choice of the FFT signal for the feature extraction is motivated by the necessity to extract relevant information from the transient signal $z_r(t)$. The FFT is able to extract the magnitude of each frequency in the response signal $z_t(t)$. It is assumed that each grounding configuration has its own FFT pattern, even in different soils, i.e. the objective is to extract the rods contributions to the response.

4) Pattern Classification and Model Learning Systems: There are different ways to compose the model $h(\cdot)$ by using the pattern recognition framework. However, the models used in this paper were limited to the state of the art in classification models, the Adaptive Boosting (Adaboost), Support Vector Machine (SVM) and Random Forest Classifiers (RF) classifiers. Each of these models are data-driven models, i.e. they learn the parameter $\boldsymbol{\zeta}$ trough examples. The objective in this paper is to make the Adaboost, SVM and RF classifiers learn different FFT patterns for different GS configurations. Each of the pattern classification models are described as follows.

a) Adaboost: The AdaBoost (Adaptive Boosting) model, proposed in [11], is a pattern recognition model algorithm that combines weak classifiers to form a strong classifier. The definition of weak and strong classifiers is described in the probably approximately correct (PAC) framework [12]. The output $h(\cdot)$ of Adaboost is given by

$$h(\mathbf{x}_i) = \operatorname{sign}\left(\sum_{k=1}^{T} a_k h_k(\mathbf{x}_i, \boldsymbol{\theta}_k)\right), \tag{6}$$

where h_k are the weak classifiers, T are the number of weak classifiers, and $\boldsymbol{\zeta} = \{a_k, \boldsymbol{\theta}_k\}_{k=1}^{T}$, are the set of parameters of the Adaboost model. After defining the number of weak classifiers, the learning step is responsible to tune the parameters $\boldsymbol{\zeta}$. There are plenty of algorithms used to learn the parameters $\boldsymbol{\zeta}$, the classical one is given in [11]. Recently, an approach based on the gradient descent algorithm was proposed in [13]. The Adaboost is primary designed to deal with binary classification problems, the extension of Adaboost for multiclass classification is done can be done using the max-wins rule [14].

b) Support Vector Machine: The support vector machine (SVM) model is based on the statistical learning theory, developed by [15]. It works by identifying the best separating hyperplane (the plane with maximum margins) between the two classes of the training samples within the feature space. The output $h(\cdot)$ of SVM model is given by:

$$h(\mathbf{x}_i) = \text{sign}\left(b + \sum_{\text{SV}} \lambda_i y_i K(\mathbf{x}_i, \mathbf{x}_j)\right), \tag{7}$$

where SV denotes the number of support vectors, b denotes the bias, $K(\mathbf{x}_i, \mathbf{x}_j)$ is the Kernel function and λ_i are the Lagrange multipliers, the parameters $\boldsymbol{\zeta}$ of SVM are the Lagrange multipliers λ_i. In this paper the radial basis function (RBF) Kernel is going to be employed, it is given by

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma||\mathbf{x}_i - \mathbf{x}_j||) \tag{8}$$

When using the SVM algorithm and the RBF kernel function, the parameters C and γ should be set. The learning of SVM is done by solving the quadratic programming (QP) problem described in [15]. The SVM is primary designed to deal with binary classification problems, the extension of Adaboost for multiclass classification is done can be done using the max-wins rule [14]

c) Random Forests: The Random Forests (RF), proposed by [16], can be seen as a generalization of decision trees (DT). The DT divides the feature space into sets of disjoint rectangular regions and learns a simple model for each these regions. The RF grows T DT's, where for each tree the training samples and features are randomly assigned. This strategy wants to reduce the variance of the classifier and make it more robust. The output of the RF is given by

$$h(\mathbf{x}_i) = \textit{majority vote}\,(T_k(\mathbf{x}_i))_{k=1}^{T}, \tag{9}$$

where $T_k(\mathbf{x}_i$ is the output of DT k for the input $\mathbf{x}_i$. The *majority vote* operand outputs the value that appears in the major number in $(T_k(\mathbf{x}_i))_{k=1}^{T}$. It means that means that forest chooses the classification having the most votes (over all the trees in the forest). The number T of DT should be set. Then, each of parameters $\boldsymbol{\zeta}$ of DT's can be learned by using the C4.5 learning algorithm [17].

B. Obligatory Steps

In this section the obligatory steps, necessary to compose the proposed system is going to be described. The obligatory steps make use of the shared systems to compose themselves.

1) Data Collection and Training Data Set Definition: This step is responsible for the data collection and training dataset definition. In this step the following systems are used: excitation, data acquisition and feature extraction. The data collection step is performed by defining the soils that are going to be used to extract the characteristics of different GS configurations. In this step it is necessary to assure the major variability as possible of the soil characteristics. This variability can be assured by setting different locations and weather conditions for the data collection. Then, the GS that the system is going to discriminate should be mounted in each of these different soils. Moreover, it is necessary to assure that the equipments, ground rods, the distances d_v and d_i are the same in all experiments.

2) Model Definition: This step makes use of the pattern classification and model learning system. This step is responsible to define the pattern recognition model that is going to be used to evaluate a new GS configuration. For that purpose, any model can be tuned using the training data, acquired in the data collection step. The model that gives the best classification performance should be selected as the final model.

V. Experimental Results

In order to evaluate the proposed grounding discrimination system, a set of 26 controlled experiments, in distinct soils and in different locations were conducted. In all the soils, the configuration with one, two, three and four ground rods in straight line were mounted. All the ground rods have $2.4m$ length with $0.0150m$ diameter, and for the configurations with two, three and four ground rods, the rods were spaced in intervals of $3m$. For all the 26 experiments, the excitation system was applied in the GS'sas described in Section IV-A1 and the voltage and current responses as described in Section IV-A2 were acquired through the voltage and current electrodes. Then, a total of 104 response data were acquired, where 26 belongs to the each one of the existing configurations.

As discussed in the introduction, the objective of the proposed method is to detect an incorrect installation of the GS. From the pattern recognition point of view it can be done as a multiclass classification problem, as discussed III. It has the ability to discriminate among different grounding configurations and has the capability to indicate the exact GS configuration present in the soil. The proposed method evaluates the accuracy, benefits and drawbacks of this approach in the next subsections.

A. Experimental Settings

To evaluate the proposed method in therms of its classification accuracy, the following methodology was used. From the 26 experiments, where each one is composed by the response data of the grounding system with one, two, three and four ground rods, 25 were used to train the pattern recognition model, and the remaining one, which is not part of the training set, was used to evaluate the prediction accuracy of the trained model. This was repeated such that all the soils were used as the test set, so that the information of test set is not included in the training set. This approach simulates the reality, where there is no information regarding the soil where the GS is going to be inspected. This approach is illustrated in Fig. 6.

Regarding the tunning parameters of the pattern recognition models, they were selected based on the same procedure as above; for the Adaboost the number of weak classifiers T, the margin and kernel parameters C, γ of the SVM model, and the number of DT in the RF classifiers.

The performance of each model is going be evaluated based on its rate of accuracy (ACC).

$$\text{ACC} = \frac{\sum_{i=1}^{N} I(h(\mathbf{x}_i), y_i)}{N}, \tag{10}$$

where $I(h(\mathbf{x}_i), y_i)$ is defined as:

$$I(h(\mathbf{x}_i), y_i) = \begin{cases} 1 & \text{if } h(\mathbf{x}_i) = y_i, \\ 0 & \text{if } h(\mathbf{x}_i) \neq y_i, \end{cases} \tag{11}$$

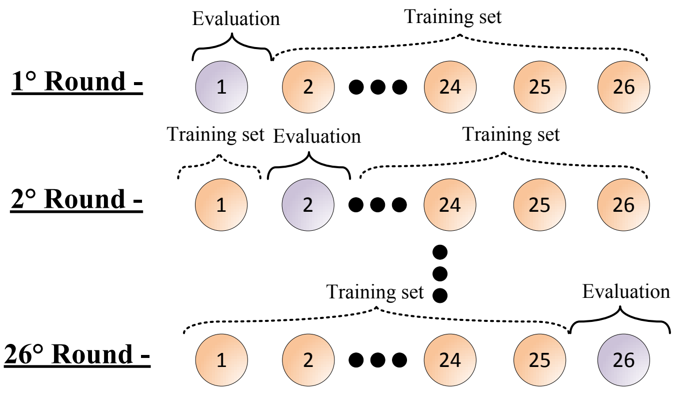


Fig. 6. Representation of the adopted training and test strategy.

The ACC measures the rate of correctness of the model and it is calculated as the sum of correct classifications divided by the total number of classifications.

B. Collected Data

A example of the collected data is shown in Fig. 7. This shows the first 250 samples of $v_r(t)$ and $i_r(t)$ signals for different grounding configurations and two distinct soils; the soils are labeled as soil #1 and soil #2. Figs. 7(a) and 7(e) show the response for the GS composed of one ground rod, Figs. 7(b) and 7(f) show the response for two ground rods, Figs. 7(c) and 7(g) show the response for three ground rods and Figs. 7(d) and 7(h) show the response for four ground rods. As can be noticed, different soils produces different responses from each other even with the same GS configuration.

C. Analysis of the Experimental Results

In the multiclass classification approach the objective is to learn several models $h(\cdot)$, to discriminate among the four classes, which correspond to the configuration with one, two, three and four ground rods, respectively. The RF, Adaboost and SVM models were trained using the strategy described previously. The results are given in therms of accuracy and also by the truth table shown in Table I. The truth table indicates the percentage of each GS configuration which was correctly assigned/classified by the classifiers (this is represented by the diagonal line), and the percentage of each class which was incorrectly assigned/classified.

By Table I, the RF model has a accuracy of 71% (mean value of the diagonal line), the best one among the two other models. In a practical perspective can identify properly if 71 out of 100 GS's are installed properly. Analyzing the truth table present in Table I, it can be seen that when the real GS is composed by one rod, the accuracy of output system is 80.0%, and in 19.2% it assigns as output two ground rods, which is the wrong output. For the case where the real GS is composed by two ground rods, the proposed system gives in 69.2% of cases two ground rods as output, and in 23.0% of the cases it gives one ground rod as output (wrong output), and in 7.7% of the cases it gives three output rods as output (again wrong output). The RF model gives the correct output in

TABLE I
TRUTH TABLE AND ACCURACY FOR RF, ADABOOST AND SVM MODELS.

RF		Output of the Proposed System			
		one rod	two rod	three rods	four rods
Real	one rod	80.8%	19.2%	0%	0%
	two rods	23.0%	69.2%	7.7%	0%
	three rods	7.7%	0%	65.4%	26.9%
	four rods	3.8%	3.8%	23.0%	69.2%
Accuracy:	71.1%				
Adaboost		**Output of the Proposed System**			
		one rod	two rod	three rods	four rods
Real	one rod	73.1%	26.9%	0%	0%
	two rods	19.2%	65.4%	7.7%	7.7%
	three rods	7.7%	7.7%	57.7%	26.9%
	four rods	0%	7.7%	26.9%	69.2%
Accuracy:	66.3%				
SVM		**Output of the Proposed System**			
		one rod	two rod	three rods	four rods
Truth	one rod	76.9%	23.1%	0%	0%
	two rods	26.9%	46.2%	19.2%	3.8%
	three rods	3.8%	23.1%	42.3%	30.8%
	four rods	3.8%	7.7%	26.9%	61.5%
Accuracy:	57.7%				

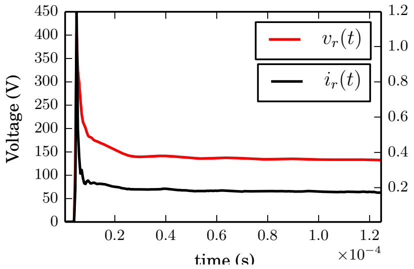
65.5% when it is composed by three ground rods, and it gives in 7.7% of the cases one ground rod as output (wrong output) and in 26.9% four ground rods as output (wrong output). For the case of four ground rods, it gives the right output in 69.2% of the cases and in 23.0% gives three ground rods as output (wrong output), and in 3.8% of the cases it gives one, and two ground rods as output (wrong output).

The Adaboost reached a accuracy of 66.3% and the SVM reached a accuracy of 57.7%. Similar conclusions of the truth table can be drawn for the Adaboost and SVM classifiers, as described in the last paragraph.

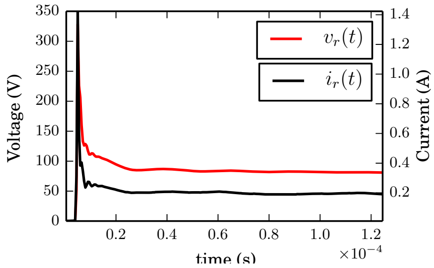
From the truth table, and for all classifiers, it is possible to conclude that all the models has a moderate rate of accuracy when discriminating GS with close number of ground rods. This is more evident for the cases of three and four rods, where the classifiers have more difficult to discriminate. From the grounding perspective, the rate of accuracy achieved, even not being so high, it is a acceptable result, since it will reduce the uncertainty of installed GS. However, from the pattern recognition perspective, the obtained rate of accuracy of the classifiers can be motivated by the fact that only few numbers of samples were used for training the models (only 25 for training and 1 for testing) and also because the presence of a large number of input features 125 in total, which makes the classification a hard task. By increasing the number of training samples, probably the number of accuracy of all classifiers would increase.

VI. CONCLUSIONS

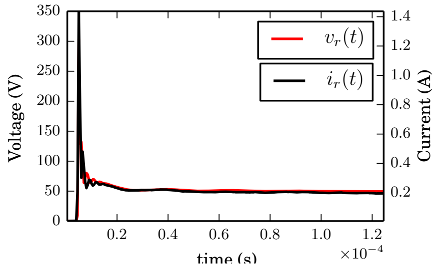
In this work a innovative method for automatic discrimination of different GS was presented. The proposed approach is based on the pattern recognition framework. The proposed system is based in four main systems: excitation, data acquisition,



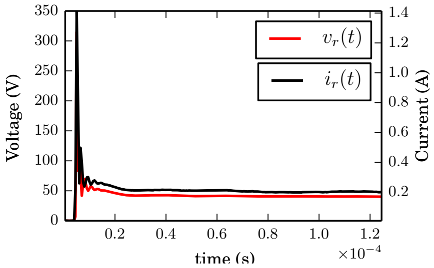
(a) $v_r(t)$ and $i_r(t)$ values for one ground rod and soil #1



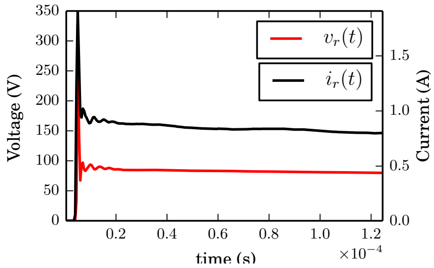
(b) $v_r(t)$ and $i_r(t)$ values for two ground rods and soil #1.



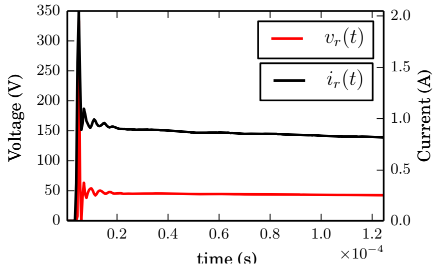
(c) $v_r(t)$ and $i_r(t)$ values for three ground rods and soil #1.



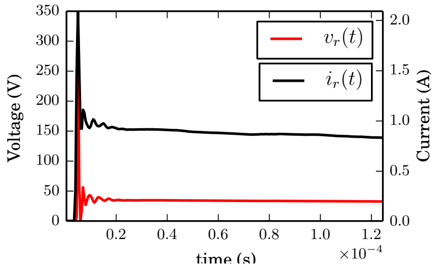
(d) $v_r(t)$ and $i_r(t)$ values for four ground rods and soil #1.



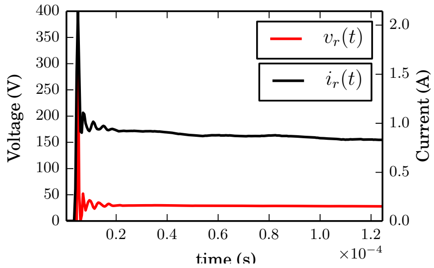
(e) $v_r(t)$ and $i_r(t)$ values for one ground rod and soil #2.



(f) $v_r(t)$ and $i_r(t)$ values for two ground rods and soil #2.



(g) $v_r(t)$ and $i_r(t)$ values for three ground rods and soil #2.



(h) $v_r(t)$ and $i_r(t)$ values for four ground rods and soil #2.

Fig. 7. First 250 samples of $v_r(t)$ and $i_r(t)$ signals, for two different soils present in the experimental part, for different GS configurations.

feature extraction and pattern classification systems. In this work each of these components were detailed described. The proposed approach allows the inspection of GS, reducing the number of wrong GS installations, increasing the confiability and reliability of the supply network.

To evaluate the proposed system a set of controlled experiments were conducted in 26 soils and GS configurations with one, two, three and four rods were mounted and their patterns were extracted by using the proposed methodology. Three classification models, RF, SVM and Adaboost models were evaluated regarding its classification accuracy. The RF model reached the highest rate of accuracy, with the value of 71%.

The proposed work demonstrates that it is possible to determine the topologies grabbed in the soil in 71% of the tested cases, even with a small number of samples (only 25 and 1 for testing) for composing the model. This rate of accuracy is very satisfactory from the GS perspective. It will reduce drastically the number of wrong GS installations, since in most of the real cases the grounding system is not inspected after it is covered. By using this method, one can reduce the uncertainty about what is the GS configuration installed.

This work has proposed an innovative work with original objectives. This innovative approach opens several gates and challenges for the grounding and pattern recognition communities, since this can be seen as a multiclass problem with few numbers of samples and many features.

Future works will address the development of strategies to increase the rate of accuracy, by improving the steps of feature extraction and classification model. Furthermore, it will also address the developments of a stackless approach, by removing the auxiliary electrodes.

References

[1] "IEEE recommended practice for grounding of industrial and commercial power systems," *IEEE Std 142-2007 (Revision of IEEE Std 142-1991)*, pp. 1–225, Nov 2007.

[2] C. Mazzetti and G. M. Veca, "Impulse behavior of ground electrodes," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-102, no. 9, pp. 3148–3156, Sept. 1983.

[3] M. Ramamoorty, M. M. B. Narayanan, S. Parameswaran, and D. Mukhedkar, "Transient performance of grounding grids," *IEEE Transactions on Power Delivery*, vol. 4, no. 4, pp. 2053–2059, Oct 1989.

[4] L. Grcev and F. Dawalibi, "An electromagnetic model for transients in grounding systems," *IEEE Transactions on Power Delivery*, vol. 5, no. 4, pp. 1773–1781, Oct 1990.

[5] A. F. Otero, J. Cidrás, and J. L. del Alamo, "Frequency-dependent grounding system calculation by means of a conventional nodal analysis technique," *IEEE Transactions on Power Delivery*, vol. 14, no. 3, pp. 873–878, Jul 1999.

[6] L. Grcev and M. Popov, "On high-frequency circuit equivalents of a vertical ground rod," *IEEE Transactions on Power Delivery*, vol. 20, no. 2, pp. 1598–1603, April 2005.

[7] R. Alipio and S. Visacro, "Impulse efficiency of grounding electrodes: Effect of frequency-dependent soil parameters," *IEEE Transactions on Power Delivery*, vol. 29, no. 2, pp. 716–723, April 2014.

[8] ——, "Modeling the frequency dependence of electrical parameters of soil," *IEEE Transactions on Electromagnetic Compatibility*, vol. PP, no. 99, pp. 1–9, 2014.

[9] A. D. Conti and S. Visacro, "A simplified model to represent typical grounding configurations applied in medium-voltage and low-voltage distribution lines," in *IX International Symposium on Lightning Protection*, November 2007, pp. 1–6.

[10] M. S. Naidu and V. Kamaraju, *High Voltage Enginnering*, 2009.

[11] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119 – 139, 1997.

[12] L. G. Valiant, "A theory of the learnable," *Commun. ACM*, vol. 27, no. 11, pp. 1134–1142, 1984.

[13] B. Babenko, M.-H. Yang, and S. Belongie, "A family of online boosting algorithms," in *IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops) 2009*, September 2009, pp. 1346–1353.

[14] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc. 2001.

[15] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 1999.

[16] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[17] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. 1993.