Journal of Control, Automation and Electrical Systems (2018) 29:640–649
https://doi.org/10.1007/s40313-018-0406-7

# Fault Location in the Smart Grids Context Based on an Evolutionary Algorithm

Danilo de Souza Pereira[1] · Carlos Frederico Meschini Almeida[1] · Nelson Kagan[1]

Received: 20 December 2017 / Revised: 13 June 2018 / Accepted: 8 August 2018 / Published online: 14 August 2018


**Abstract**
Advances in communications technologies, data processing and storage benefit power distribution utilities, allowing them to enhance the use of data provided by field monitors, remote-controlled switches and smart meters. Today, utilities can gather a variety of data regarding the power grid operation, such as customers demands, alarms and measurements, taking steps toward the Smart Grids. An interoperability bus (IB) can provide those data to any other corporate system, allowing one to develop power grid operational tools executed by distribution management systems (DMS). In this context, the present paper proposes a new fault location methodology for real power distribution networks, that resorts to data provided by an IB, such as alarms from protection relays and fault indicator sensors, and measurements from power quality monitors and smart meters. The methodology can be implemented in the DMS level and is based on evolutionary strategies, which is responsible for estimating the exact location of faults in the MV level of power distribution networks. The effect of the availability of data as alarms and measurements is assessed, considering a real 319-km-long power distribution feeder. Test results obtained from 113 short-circuit cases have indicated that the locating error is inferior to 2.9%.

**Keywords** Fault location · Smart grids · Power quality · Distribution automation

## 1 Introduction

### 1.1 Faults in Distribution Systems

Overhead electric power distribution systems are usually subject to animal and vegetation contact, vehicle collisions, bad weather conditions. The resulting short-circuit condition is extinguished by protection devices operation, which affects many customers with power outage. It causes discomfort, economic and material losses (Cordova and Faruque 2015). Such reality demands tools and methodologies for efficient and effective FL, fault isolation, repair and service restoration, in order to improve the PQ.

✉ Danilo de Souza Pereira
danilopereira@usp.br

Carlos Frederico Meschini Almeida
cfmalmeida@usp.br

Nelson Kagan
nelsonk@pea.usp.br

[1] Universidade de São Paulo, São Paulo, Brazil

### 1.2 Fault Location in Traditional Power Distribution Grids

Initial frameworks designed for FL addressed the location of electric faults in transmission systems (Saha et al. 2002). For instance, Girgis and Fallon (1992) proposes the use of DFR (Digital Fault Recorders) to register voltages and currents waveforms at the beginning of a power distribution feeder. The FL methodology described in Zhu et al. (1997) aims at distribution systems and employs similar mathematical approach as Girgis and Fallon (1992), but considering a database for the grid modeling and iteratively computing the fault distance for each section of a power distribution feeder. Such approach overcomes the problem of phase imbalance; however, it may lead to multiple possible locations, when only the measurements from one position in the power grid is available during the occurrence of a fault.

Based on Zhu et al. (1997), the authors of reference Senger et al. (2005) have proposed a comprehensive framework for FL in distribution systems, based on waveforms recorded by protection IEDs allocated at the substation for each power distribution feeder. This work brought contributions to the subject as it automates the FL procedure, proposing an opti-

mal data flow, accurate system modeling and simple firmware updates. On the other hand, the system does not update the network model (such as switches states) automatically. Then, it may not match the real power grid states and lead to wrong results.

Different Artificial Intelligence (AI) techniques for FL have been considered such as Artificial Neural Networks (ANN), Fuzzy Logic, Genetic Algorithms, among others. Alshaher et al. (2003) proposes the use of ANN for FL, which should be previously trained with some of the power grid variables. In Farias et al. (2016), however, the authors propose a FL technique attempting to avoid time consuming training sessions, with on-line training. In Guerra and Kagan (2009), the authors propose a FL methodology based on voltage and current quantities recorded by PQ monitors deployed along a power transmission system. Those distributed measurements are integrated in an Evolutionary Strategy.

The authors of Manassero et al. (2017) propose a FL methodology based on heuristics. A Patter Search method implementation estimates the fault distance and type. Among the inputs, however, measurements at the power feeder beginning are the only updated information. Considering fixed switches states may yield to wrong fault location results.

### 1.3 Fault Location in Smart Grids

Technical operation levels of power distribution network establish losses reduction, voltage levels improvement and increase in grid availability and reliability. Remote switching and monitoring allow decision-making for an efficient grid operation. Capacitor banks, on-load tap changers and voltage regulators provide flexible operation as it automates voltage and reactive power adjustments remotely and on real-time. Smart Meters connected through communication systems enable remote metering, checking meter operational status and notifying power outages. The deployment of intelligent devices throughout the distribution power grid has allowed utilities to accurately measure its electric quantities in real time.

The aforementioned Smart Grid environment represents a real opportunity for the development of tools such as fault location algorithms. Some works have contributed toward this research field. In Trindade et al. (2014), the authors propose the use of Smart Meters for FL, based on voltage sags recording, which is a functionality available in some meters. Such proposal assumes a fully functional communication infrastructure, which may limit the use of FL algorithms to situations where there is no loss on the data transmitted. Another drawback is the heavy data flow caused by all the affected customers.

Toward the use of sensors, Džafić et al. (2018) presents a FL methodology based on telemetered Fault Indicators (FI) deployed along the power feeder. However, many FI devices are needed and a heavy data flow is caused. Moreover, topological data are not updated in real time, which may cause the fault location results to be inaccurate. Reference Parker and McCollough (2011) indicates that the utilization of current sensors throughout power distribution feeders improves the FL accuracy and yields a more efficient operation of the distribution system.

The herein paper proposes an innovative FL methodology, addressing the problem through data integration in the a Smart Grid context. Its main innovation consists of locating faults based on real-time information, provided by a comprehensive set of field equipment and corporate systems, supported by data integration. Considering the data provided by such field equipment, the paper quantifies how one may locate faults more accurately in real power distribution networks through a meta-heuristic approach.

In terms of IT systems, Supervisory Control and Data Acquisition System (SCADA) and protection systems are responsible for gathering alarms and real-time measurements from field equipment, such as protection IEDs, fault indicators and smart meters. The alarms work as the FL trigger, whereas the real-time measurements are the main FL input, supporting the search for the fault position.

Systems focused on database, such as Meter Data Management (MDM) and Geographic Information System (GIS), gather detailed and accurate power grid modeling information, providing updated power grid topology, customers demands and switches statuses. That information increases the FL algorithm accuracy, as the digital model attempts to match the real distribution power grid.

## 2 Methodology

### 2.1 Assumptions

The FL methodology locates faults only in Medium Voltage (MV) level and considers all the Low Voltage (LV) loads connected to the MV/LV transformer's secondary terminal. Distribution utilities continuity indexes (SAIDI and SAIFI, for instance) are mostly impacted by faults in MV level.

For the FL algorithm, the methodology works with load blocks. Load blocks are portions of the power distribution grid delimited by NO or NC switches, such as fuses, reclosers, secctionalizers, regular switches, etc. Thus, a load block groups line sections, buses, MV/LV transformers, loads, capacitor banks, voltage regulators, etc. This grouping process facilitates the search for the fault location in power distribution grids with many branches. Concerning the smart meters power outage alarms, the FL methodology associates the set of LV customers alarms with a unique alarm attached to the transformer MV bus.

The fault indicators provide the fault event alarm and the fault current magnitude. The PQ monitors measure voltages and currents magnitudes, which may also be measured by protection relays.

## 2.2 FL Process

To accomplish the FL based on up-to-date information from the utility corporate systems, it is necessary to follow the steps presented by Fig. 1. The FL process is initiated by the arrival of data regarding alarms from corporate systems. Then, it allows a first search area restriction. This is followed by the fault type determination, through a phase currents analysis. An Evolutionary Strategy (ES) algorithm provides an accurate determination of the fault characteristics (fault resistance and location).

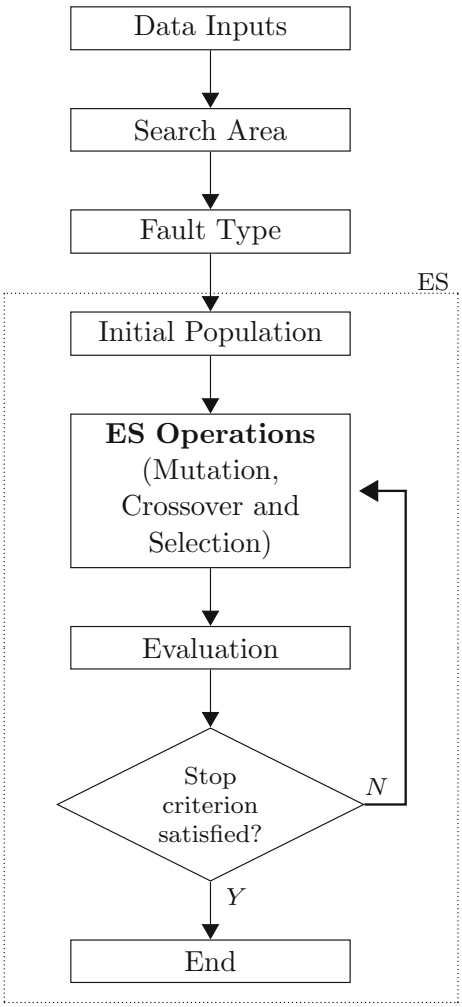


**Fig. 1** Fault location overall process

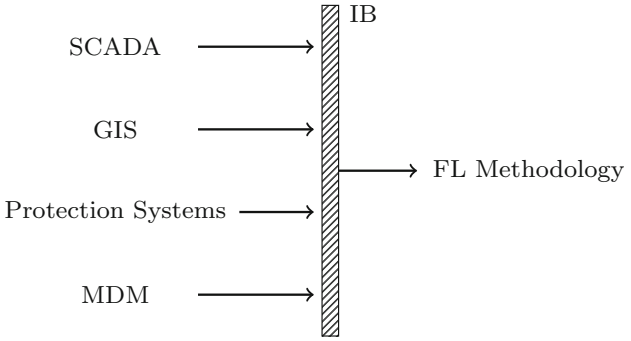


**Fig. 2** Data flow supporting the FL methodology

## 2.3 Data Inputs

The utility corporate systems gather a great variety of the distribution power grid data. These systems are supposed to support the FL methodology with the following data:

- *SCADA* Up-to-date information about field equipment status, such as remotely controlled switches, current sensors (currents recordings) and PQ monitors (voltages and currents recordings);
- *GIS* Customers georeferenced data, supporting digital modeling of the power grid;
- *Protection systems data* Protection relays parameters and waveforms records;
- *MDM* Customers readouts and power outages alarms;

An IB should integrate those systems data with the FL methodology. Figure 2 illustrates the resulting data flow.

The FL application is intended to be executed in a distribution management system (DMS), aimed at dealing with complex algorithms for grid operations optimization and decision-making support. The FL is based on relays waveforms, alarms from MV current sensors and protection schemes tripping alarms. The SCADA system is likely to contain those data.

The GIS system is responsible for providing data to build an up-to-date power grid digital model. The FL application also utilizes smart meters power outage alarms, which can be obtained from the MDM system. The methodology considers an integrated environment, in such a way that data may be exchanged seamlessly.

## 2.4 Defining the Search Area

The search area determination is based on incoming data, such as power outage alarms, alarmed fault current sensors and tripped remotely controlled protection devices. The analysis of those data may lead to two main situations, as follows.

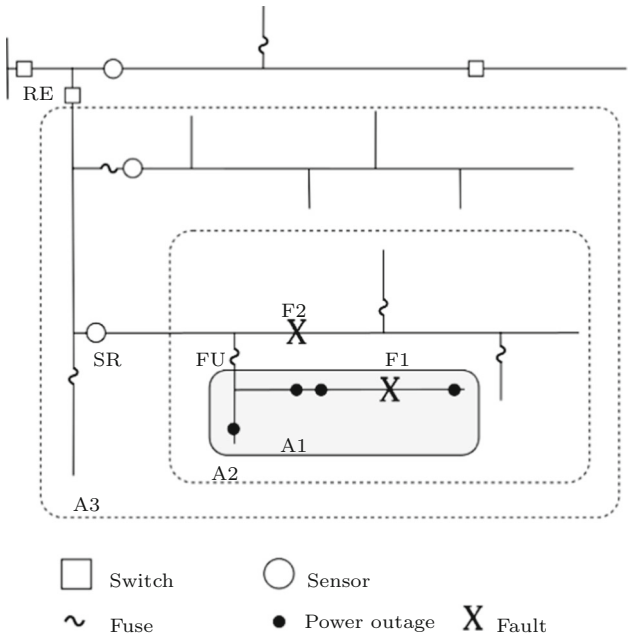


**Fig. 3** Situations in determining the search area

- *Situation 1* The fault happens downstream a fuse switch. When this occurs, no remotely controlled protection device is alarmed. In the Fig. 3 scheme, this situation happens with fault **F1**. The smart meters (small black dots) of the affected customers detect the fault event and send out power outage alarms. The Sensor *SR* is the most downstream to be alarmed by the fault current. Fuse switch *FU* trips and no remotely controlled protection operates. The smart meters power outage alarms information yields to search area **A1**. Meanwhile, the sensor information yields to search area **A2**. Area **A1** is considered the search area, because it is smaller than **A2**.
- *Situation 2* The fault happens downstream a recloser switch and upstream the fuses, causing the former to trip. This tripping event is sent to SCADA. Sensors information may also support the search area determination. In the Fig. 3 scheme, this situation occurs with fault **F2**. The most downstream alarmed sensor is *SR*, which yields to a candidate search area **A2**. Meanwhile, the recloser *RE* tripping operation leads to a candidate search area **A3**. Area **A2** is considered the search area because it is smaller than **A3**.

The search area carries out a division process, in such a way that each line section is represented by a real interval $[L_{\%}^{(1)}, L_{\%}^{(2)}]$, belonging to [0%, 100%]. Such process is necessary because it allows the algorithm to refer to each particular line section.

For a line section $i$, with length $L_i$, the calculation of $L_{\%_i}^{(1)}$ and $L_{\%_i}^{(2)}$ is based on the accumulated line lengths ($L_i^{\text{accum}}$),

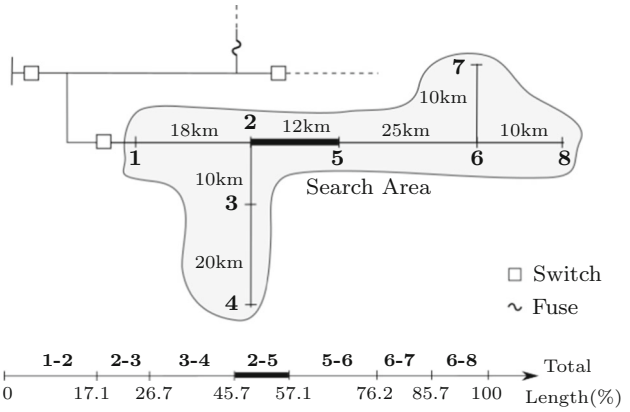


**Fig. 4** Dividing the search area

**Table 1** Line segments division

| Segment $i$ | $L_i$ (km) | $L_i^{\text{accum}}$ | $L_{\%_i}^{(1)}$ | $L_{\%_i}^{(2)}$ |
|---|---|---|---|---|
| 1–2 | 18 | 18 | 0 | 17.1 |
| 2–3 | 10 | 28 | 17.1 | 26.7 |
| 3–4 | 20 | 48 | 26.7 | 45.7 |
| 2–5 | 12 | 60 | 45.7 | 57.1 |
| 5–6 | 25 | 85 | 57.1 | 80.9 |
| 6–7 | 10 | 95 | 80.9 | 90.5 |
| 6–8 | 10 | 105 | 90.5 | 100 |

as detailed by Eqs. 1, 2 and 3, where $L_{\text{total}}$ is the distribution feeder total length.

$$L_i^{\text{accum}} = \sum_{k=1}^{i-1} L_k \tag{1}$$

$$L_{\%_i}^{(1)} = 100 \cdot L_i^{\text{accum}} / L_{\text{total}} \tag{2}$$

$$L_{\%_i}^{(2)} = 100 \cdot \left(L_i^{\text{accum}} + L_i\right) / L_{\text{total}} \tag{3}$$

For the line sections of the distribution power grid in Fig. 4, the line segments division is presented by Table 1.

## 2.5 Fault Type Determination

Prior to the FL process, the FL application should estimate the fault type, determining the affected phases by the fault event. An assumption is made: only MV faults are considered and there are no line transformers along the feeder.

In order to proceed the fault type determination, some quantities are defined and computed, based on the phase currents measurements during the fault. Equations 4 and 5 define the average current $I_{\text{av}}$ and the phase currents deviations $\text{d}I_{\text{ph}}$, with ph = $a$, $b$ or $c$.

$$I_{\text{av}} = \frac{|\dot{I}_a^{\text{fault}}| + |\dot{I}_b^{\text{fault}}| + |\dot{I}_c^{fault}|}{3} \tag{4}$$

$$\text{d}I_{\text{ph}} = \left( \frac{|\dot{I}_{\text{ph}}^{\text{fault}}| - I_{av}}{I_{av}} \right) \cdot 100(\%) \tag{5}$$

The possible outputs for the fault classification considered in this paper are: ABC (three-phase fault), AG (phase A to ground fault), ABG (phase A and phase B to ground fault) and AB (phase A to phase B fault). The fault classification logical reasoning considered is:

- Three-phase fault (3Ph):
  **If** ($|\text{d}I_a| < \text{max1}$ **And** $|\text{d}I_b| < \text{max1}$ **And** $|\text{d}I_c| < \text{max1}$) **Then**: ABC.
- Single-phase fault (1Ph):
  **If** ($\text{d}I_a > 0$ **And** $|\text{d}I_a| > \text{min1}$ **And** $\text{d}I_b < 0$ **And** $\text{d}I_c < 0$) **Then**: AG.
  **Else, if** ($\text{d}I_b > 0$ **And** $|\text{d}I_b| > \text{min1}$ **And** $\text{d}I_a < 0$ **And** $\text{d}I_c < 0$) **Then**: BG.
  **Else, if** ($\text{d}I_c > 0$ **And** $|\text{d}I_c| > \text{min1}$ **And** $\text{d}I_a < 0$ **And** $\text{d}I_b < 0$) **Then**: CG.
- Phase-to-phase fault (2Ph or 2Phg):
  **If** ($\text{d}I_a > 0$ **And** $\text{d}I_b > 0$ **And** $\text{d}I_c < 0$) **Then**:
  **If** $|I_0/I_1| > \text{min2}$ **Then**: ABG.
  **Else, if** $|I_0/I_1| \leq \text{min2}$ **Then**: AB.

  **If** ($\text{d}I_b > 0$ **And** $\text{d}I_c > 0$ **And** $\text{d}I_a < 0$) **Then**:
  **If** $|I_0/I_1| > \text{min2}$ **Then**: BCG.
  **Else, if** $|I_0/I_1| \leq \text{min2}$ **Then**: BC.

  **If** ($\text{d}I_c > 0$ **And** $\text{d}I_a > 0$ **And** $\text{d}I_b < 0$) **Then**:
  **If** $|I_0/I_1| > \text{min2}$ **Then**: ACG.
  **Else, if** $|I_0/I_1| \leq \text{min2}$ **Then**: AC.

where

- max1 Maximum deviation of a phase current ($\text{d}I_{\text{ph}}$) for a three-phase fault.
- min1 Minimum deviation of a phase current ($\text{d}I_{\text{ph}}$) for a phase-to-ground fault.
- min2 Minimum rate of the zero sequence fault current with respect to the positive sequence fault current ($I_0/I_1$) for a double phase-to-ground fault.

The parameters max1, min1 and min2 have to be determined empirically, based on historical records fault currents measurements. Those parameters should be unique for a particular power feeder.

### 2.6 Fault Location Determination with Evolutionary Strategy

The fault location determination is effectively accomplished with an ES algorithm, (details regarding ES algorithm are available in "Appendix"). The codification of the individuals considers two parameter vectors ($x$, $R_{\text{f}}$) and their standard deviations ($\sigma_x, \sigma_{R_{\text{f}}}$). A particular individual represents a solution for the FL problem, with a fault location $x$, in the range [0, 100], and a fault resistance $R_{\text{f}}$ in the range $[0, R_{\text{f}_{\max}}]$, with $R_{\text{f}_{\max}}$ the maximum considered fault resistance.

An individual (representing a particular fault) is evaluated through a fault simulation, using OpenDSS software. The resulting quantities (calculated voltages and currents during the fault) are considered by the FL algorithm to calculate errors between the fault simulation and the fault event.

The individual's evaluation leads to the error indices $\epsilon_{\text{V}}$ (for voltages errors) and $\epsilon_{\text{I}}$ (for current errors), which may be calculated through Eqs. 6 and 7. In those equations, $N_{\text{V}}$ and $N_{\text{I}}$ are the numbers of voltage and current measurements available, respectively. $|\dot{V}_j^{\text{comp}}|$ and $|\dot{I}_j^{\text{comp}}|$ are the calculated values for the voltage and current magnitudes at the measuring points and $|\dot{V}_j^{\text{meas}}|$ and $|\dot{I}_j^{\text{meas}}|$ are the voltage and current magnitudes actually measured. Such values are normalized by the rated phase voltage $V_{\text{rated}}$ and the rated current $I_{\text{rated}}$, respectively. Based on $\epsilon_{\text{V}}$ and $\epsilon_{\text{I}}$, the individual's error is computed as defined by Eq. 8, where $K_{\text{V}}$ and $K_{\text{I}}$ are the voltage and current weights, which may vary from 0 to 1 each and their sum must result in 1.

$$\epsilon_{\text{V}} = \left( \sqrt{\frac{1}{N_{\text{V}}} \sum\nolimits_{j=1}^{N_{\text{V}}} \left( \frac{|\dot{V}_j^{\text{comp}}| - |\dot{V}_j^{\text{meas}}|}{V_{\text{rated}}} \right)^2} \right) \tag{6}$$

$$\epsilon_{\text{I}} = \left( \sqrt{\frac{1}{N_{\text{I}}} \sum\nolimits_{j=1}^{N_{\text{I}}} \left( \frac{|\dot{I}_j^{\text{comp}}| - |\dot{I}_j^{\text{meas}}|}{I_{\text{rated}}} \right)^2} \right) \tag{7}$$

$$\epsilon = \frac{K_{\text{V}} \cdot \epsilon_{\text{V}} + K_{\text{I}} \cdot \epsilon_{\text{I}}}{K_{\text{V}} + K_{\text{I}}} \tag{8}$$

Ultimately, the individuals evaluation function $f_{\text{eval}}$ is defined as $f_{\text{eval}} = 1 - \epsilon$.

## 3 Tests and Results

### 3.1 Tests Description

An application supported by OpenDSS was developed to simulate fault events. The user enters the fault data (type, fault resistance and location) and the application returns measured voltages and currents and customer power outage, which are

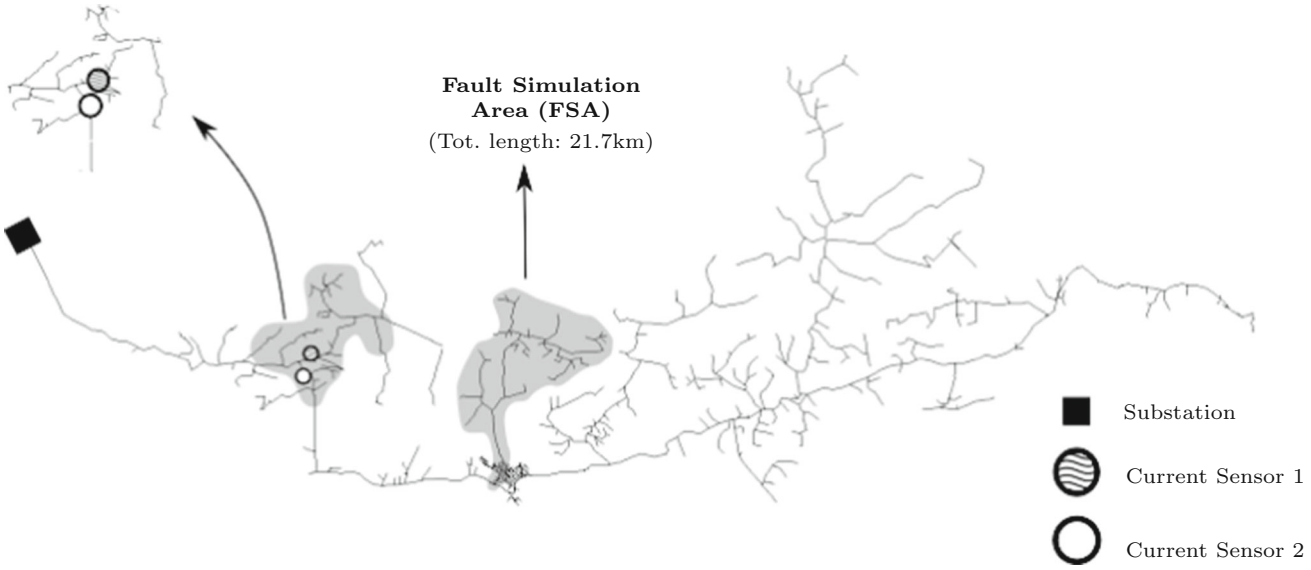


**Fig. 5** Power distribution grid and some of the simulated short-circuit points

forwarded to the FL application, that attempts to locate the fault.

A Brazilian MV power distribution feeder, shown by Fig. 5, was utilized for the simulations. Its maximum extension from the substation is 40.4 km, and summing all its line sections lengths totals 319 km. The distribution feeder has three voltage regulators equally spaced along its length, but they were disabled for the tests. There are 651 MV/LV transformers and one substation transformer.

For each of the fault types—**abc** (Three-phase), **abg** (Phase–phase–ground), **ab** (Phase–phase), and **a** (Phase–ground)—a number of faults were simulated, with fault resistance and location manually chosen in the fault simulation area (FSA). It is a 21.7-kilometer-long area, highlighted in the mentioned figure. The faults simulations are numbered from 1 to 113 and the actual location of some of them are shown in Fig. 6.

The FSA was selected because it gathers interesting characteristics in terms of topology and field equipment. First of all, it is started with a recloser switch. Its remotely controlled relay may provide electric quantities readouts during the fault. As shown by Fig. 6, its beginning has few laterals, but its final part is highly branched. The FSA topology allows testing the contribution of monitoring equipment, such as PQ monitors 1 and 2. Those equipment may distinguish, for instance, faults 1 and 11 (see Fig. 6).

Three scenarios were considered for the conducted simulations. The electrical quantities are measured during the short-circuit events and the mentioned monitoring equipment are depicted in Figs. 5 and 6.

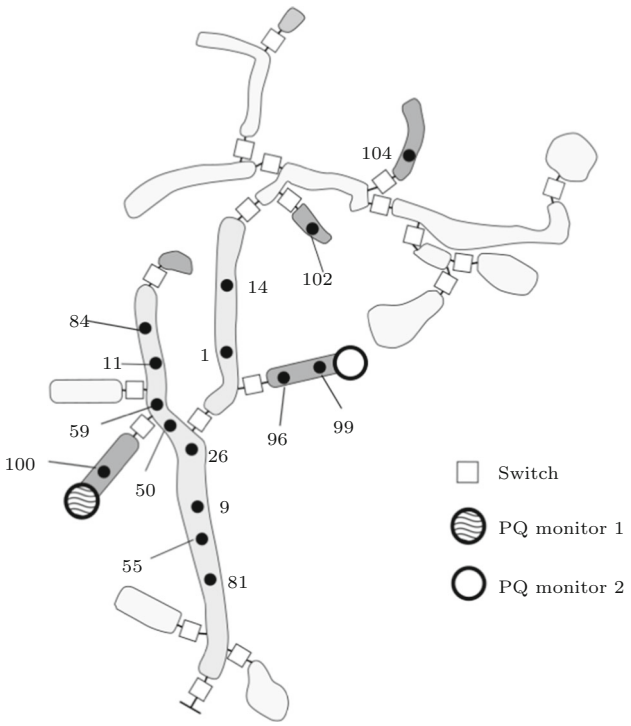


**Fig. 6** Power distribution grid and some of the simulated short-circuit points

– *Scenario 1* Circuit breaker relay measures voltages and currents and sensors 1 and 2 measure currents (see Fig. 5).
– *Scenario 2* The PQ Monitors 1 and 2 (see Fig. 6) voltages measurements are considered, along with those of Scenario 1.
– *Scenario 3* The faults are applied to the dark gray region of the fault simulation area (see Fig. 6). The affected

**Table 2** ES settings for the conducted simulations

| Parameter | Value |
|---|---|
| No. of initial individuals | 40 |
| Max. No. of individual per generation | 10 |
| Mutation probability | 0.8 |
| Descendants per individual in mutation | 6 |
| Crossover probability | 0.5 |
| Selection affected individuals | Parents and children |
| Selection criterion | Best individuals |
| Initial $\sigma_x$ ($x$ SD) | 15 m |
| Initial $\sigma_{R_f}$ ($R_f$ SD) | 10 Ω |
| Maximum $R_f$ | 15 Ω |

smart meters send out power outage alarms, which are considered along with the data of Scenarios 1 and 2.

The ES settings for the conducted simulations are listed in Table 2.

### 3.2 Results by Fault Scenario

For Scenarios 1, 2 and 3, the corresponding tables (Tables 3, 4 and 5, respectively) present the simulated faults settings and its corresponding locating errors, which are given in meters and percentage of the search area total length (21.7 km). In Fig. 6, the black dots represent some of the faults actual locations.

Calculating the average and standard deviation values of $X$ and $R_f$, one may notice that Scenario 2 had improvements with respect to Scenario 1: the average location error dropped from 629.76 m to 180.59 m, a decrease of 71.32%, due to the insertion of PQ Monitors 1 and 2. As expected, measuring voltage at different points captures the global effects of short circuits in those locations. This approach reduces the number of possible solutions, increasing the FL methodology accuracy.

Scenario 3 presented improvements with respect to Scenario 2: There was reduction in the average location error, that dropped from 180.59 m to 63.75 m, i.e., a further reduction of 64.70%. The improvement was provided by the utilization of customers power outage information, which may further restrict the search area. A smart meter outage alarm indicates that the customer belongs to the faulted line section or to its surrounding areas. Moreover, in cases that only a fuse switch operates to clear the fault, outage alarms may trigger the FL methodology and the line sections downstream the fuse is considered the search area.

Another important aspect is the drop of maximum location error, dropping from 2732 m (Scenario 1) to 1120.62 m (Scenario 2), reduction of 59%. That maximum error also

**Table 3** FL Results—Scenario 1

| Simulated faults | | | Fault location results | |
|---|---|---|---|---|
| No. | Fault type | $R_f$ (Ω) | Error (m) | Error (%) |
| 1 | abc | 0 | 167.44 | 0.77 |
| 2 | | 0 | 1793 | 8.26 |
| 3 | | 0 | 220.51 | 1.02 |
| 4 | | 0 | 1056 | 4.87 |
| 5 | | 0 | 275.71 | 1.27 |
| 6 | | 0 | 2732 | 12.59 |
| 7 | | 0 | 160.03 | 0.74 |
| 8 | | 0 | 285.31 | 1.31 |
| 9 | | 0 | 12.38 | 0.06 |
| 10 | | 0 | 13.45 | 0.06 |
| 11 | abg | 2.75 | 1717.5 | 7.91 |
| 12 | | 1.9 | 1033.24 | 4.76 |
| 13 | | 0.75 | 1043.5 | 4.81 |
| 14 | | 1.42 | 1589.2 | 7.32 |
| 15 | | 7.2 | 1374.64 | 6.33 |
| 16 | | 0.01 | 125.78 | 0.58 |
| 17 | | 2 | 128.2 | 0.59 |
| 18 | | 4.5 | 112.14 | 0.52 |
| 19 | | 1.3 | 1472 | 6.78 |
| 20 | | 1.25 | 453.4 | 2.09 |
| 21 | | 0.25 | 85.92 | 0.40 |
| 22 | | 0.25 | 186.88 | 0.86 |
| 23 | ab | 0.5 | 12.75 | 0.06 |
| 24 | | 0.01 | 436.07 | 2.01 |
| 25 | | 0.01 | 2.18 | 0.01 |
| 26 | | 0.01 | 63.07 | 0.29 |
| 27 | | 1.5 | 1813.19 | 8.36 |
| 28 | | 1 | 1581.31 | 7.29 |
| 29 | | 0.01 | 6.63 | 0.03 |
| 30 | | 0.01 | 59.23 | 0.27 |
| 31 | | 1.45 | 97.17 | 0.45 |
| 32 | | 0.01 | 85.46 | 0.39 |
| 33 | | 0.01 | 6.59 | 0.03 |
| 34 | | 0.7 | 1228.8 | 5.66 |
| 35 | a | 0.01 | 1890.65 | 8.71 |
| 36 | | 2.5 | 175.44 | 0.81 |
| 37 | | 1.25 | 19.63 | 0.09 |
| 38 | | 3.75 | 126.19 | 0.58 |
| 39 | | 0.5 | 1554.29 | 7.16 |
| 40 | | 0.75 | 2167.85 | 9.99 |
| 41 | | 0.45 | 1507.3 | 6.95 |
| 42 | | 1.5 | 126.41 | 0.58 |
| 43 | | 3.2 | 228.2 | 1.05 |
| 44 | | 7.5 | 102.65 | 0.47 |
| 45 | | 12 | 1.9 | 0.01 |
| 46 | | 5.3 | 242.18 | 1.12 |

**Table 3** continued

| Simulated faults | | | Fault location results | |
|---|---|---|---|---|
| No. | Fault type | $R_f(\Omega)$ | Error (m) | Error (%) |
| 47 | | 1.21 | 25.2 | 0.12 |
| Maximum | | | 2732.00 | 12.59 |
| Minimum | | | 1.90 | 0.01 |
| Average | | | 629.76 | 2.90 |
| SD | | | 744.55 | 3.43 |

**Table 4** FL results—Scenario 2

| Simulated faults | | | Fault location results | |
|---|---|---|---|---|
| No. | Fault type | $R_f(\Omega)$ | Error (m) | Error (%) |
| 48 | abc | 0 | 13.32 | 0.06 |
| 49 | | 0 | 10.38 | 0.05 |
| 50 | | 0 | 9.61 | 0.04 |
| 51 | | 0 | 27.06 | 0.12 |
| 52 | | 0 | 9.98 | 0.05 |
| 53 | | 0 | 16.16 | 0.07 |
| 54 | | 0 | 15.44 | 0.07 |
| 55 | | 0 | 10.1 | 0.05 |
| 56 | | 0 | 3.15 | 0.01 |
| 57 | | 0 | 9.57 | 0.04 |
| 58 | abg | 2.75 | 394.94 | 1.82 |
| 59 | | 1.9 | 178.79 | 0.82 |
| 60 | | 0.75 | 216.84 | 1.00 |
| 61 | | 1.42 | 174 | 0.80 |
| 62 | | 7.2 | 35.3 | 0.16 |
| 63 | | 0.01 | 309.32 | 1.43 |
| 64 | | 2 | 4.83 | 0.02 |
| 65 | | 4.5 | 231.04 | 1.06 |
| 66 | | 1.3 | 34.33 | 0.16 |
| 67 | | 1.25 | 193.09 | 0.89 |
| 68 | | 0.25 | 101.51 | 0.47 |
| 69 | | 0.25 | 307.43 | 1.42 |
| 70 | ab | 0.5 | 50.15 | 0.23 |
| 71 | | 0.01 | 283.59 | 1.31 |
| 72 | | 0.01 | 4.44 | 0.02 |
| 73 | | 0.01 | 16.74 | 0.08 |
| 74 | | 1.5 | 1120.62 | 5.16 |
| 75 | | 1 | 94.53 | 0.44 |
| 76 | | 0.01 | 103.57 | 0.48 |
| 77 | | 0.01 | 16.91 | 0.08 |
| 78 | | 1.45 | 53.61 | 0.25 |
| 79 | | 0.01 | 40.93 | 0.19 |
| 80 | | 0.01 | 50.66 | 0.23 |
| 81 | | 0.7 | 1059 | 4.88 |
| 82 | a | 0.01 | 445.32 | 2.05 |

**Table 4** continued

| Simulated faults | | | Fault location results | |
|---|---|---|---|---|
| No. | Fault type | $R_f(\Omega)$ | Error (m) | Error (%) |
| 83 | | 2.5 | 110.94 | 0.51 |
| 84 | | 1.25 | 588.63 | 2.71 |
| 85 | | 3.75 | 26.8 | 0.12 |
| 86 | | 0.5 | 926.18 | 4.27 |
| 87 | | 0.75 | 118.37 | 0.55 |
| 88 | | 0.45 | 283.09 | 1.30 |
| 89 | | 1.5 | 194.16 | 0.89 |
| 90 | | 3.2 | 424.1 | 1.95 |
| 91 | | 7.5 | 66.45 | 0.31 |
| 92 | | 12 | 0.13 | 0.00 |
| 93 | | 1.21 | 70.56 | 0.33 |
| 94 | | 1.21 | 32.21 | 0.15 |
| Maximum | | | 1120.62 | 5.16 |
| Minimum | | | 0.13 | 0 |
| Average | | | 180.59 | 0.83 |
| SD | | | 263.09 | 1.21 |

**Table 5** FL results—Scenario 3

| Simulated faults | | | Fault location results | |
|---|---|---|---|---|
| No. | Fault type | $R_f(\Omega)$ | Error (m) | Error (%) |
| 95 | abc | 0 | 6.13 | 0.03 |
| 96 | | 0 | 0.24 | 0.00 |
| 97 | | 0 | 1.19 | 0.01 |
| 98 | | 0 | 28.63 | 0.13 |
| 99 | | 0 | 0.33 | 0.00 |
| 100 | abg | 0.5 | 61.39 | 0.28 |
| 101 | | 0.5 | 0.53 | 0.00 |
| 102 | | 2 | 50.71 | 0.23 |
| 103 | | 1.5 | 45.34 | 0.21 |
| 104 | | 1.5 | 398.07 | 1.83 |
| 105 | ab | 0.1 | 118.09 | 0.54 |
| 106 | | 0.01 | 55.79 | 0.26 |
| 107 | | 0.5 | 39.23 | 0.18 |
| 108 | | 0.75 | 76.72 | 0.35 |
| 109 | | 0.01 | 35.58 | 0.16 |
| 110 | a | 1.87 | 222.25 | 1.02 |
| 111 | | 4.65 | 0.23 | 0.00 |
| 112 | | 5.5 | 41.55 | 0.19 |
| 113 | | 3.2 | 29.22 | 0.13 |
| Maximum | | | 398.07 | 1.83 |
| Minimum | | | 0.23 | 0 |
| Average | | | 63.75 | 0.29 |
| SD | | | 94.10 | 0.43 |

decreased from 1120.62 m (Scenario 2) to 398.07 m (Scenario 3), representing a further reduction of 64.5%. Those results lead to the assumption that fault location conducted under Scenario 3 are likely to exhibit location error below 400 m, that may be considered a good FL result.

Special comment should be made with respect to the percentage values of the average FL errors. Scenario 1 presented an average location error of 629.76 m, which represents only 2.90% of the 21.7 km total length. That percentage value dropped to 0.83% for the results of Scenario 2. The average location error in Scenario 3 was even lower: 0.29%. The error percentage analysis is a more realistic manner to analyze the actual benefits of increasing monitoring for the proposed FL methodology, because it takes into consideration the global benefits of deploying a FL methodology which would support the operational activities when locating any type of fault in real power distribution networks.

In Scenario 1, some FL tests presented errors larger than 1500 m. They occurred due to the existence of branches. In these cases, some combinations of fault locations and resistances may lead to the same measurements. Then, multiple solutions may match those measurements. Inserting two voltage monitors in some of the branches allowed the multiple solutions problem to be relieved, as shown by the results of Scenario 2 simulations.

The conducted simulations allow inferring that increasing the amount of grid monitors improves the FL scheme proposed in the herein paper. In particular, the employment of PQ monitors is a key step for reducing multiple solutions, as the "Faults Area" exhibits many branches.

## 4 Conclusions

The Smart Grids should be considered the most appropriate environment for the development of operational tools, such as FL algorithms. Such environment provides utilities with multiple data about the various power grid monitored devices, through an IB. The FL algorithm may be executed based on up-to-date power grid model, real-time alarms and field equipment measurements during the fault.

In this paper, short-circuit events were simulated considering a real power distribution feeder and the outcomes are forwarded to the devised FL application. The initial step of determining the search area significantly reduces the overall FL process, because it limits the region under investigation. The challenges of fault location in branched systems with unknown fault resistance are overcome with an Evolutionary Strategy implementation. It allows determination of both fault resistance and location, as they are assessed simultaneously.

Considering the three proposed monitoring scenarios, it was possible to infer that increasing the power grid monitoring provides benefits to the proposed FL methodology. Those improvements are even clearer as one analyses the locating errors in terms of percentage values, with respect to the entire search area. It is a more realistic manner to capture the actual benefits of increasing monitoring.

In actual utility outage management, with the increase of distribution power grid monitoring, the FL methodology should avoid multiple solutions. It means directing repairing crew to the most probable fault location, reducing the restoration time.

In cases that only a fuse switch operates to clear the fault, the operations center is not aware of the defect, because fuse switches are not remotely controlled. Then, customers smart meters outage alarms may trigger the FL methodology and the line sections downstream the fuse (usually few line sections) constitute the search area. In these cases, customers power outage information improves greatly the FL algorithm, due to the deep restriction of search area, which only contains the areas that are close to the power outage alarms notifications sources.

As field equipment provides valuable data for increasing FL accuracy, one may notice that such accuracy depends on the location of the field equipment to be installed along the power grid. Thus, future works should consider addressing the optimal allocation problem of fault sensors and PQ monitors, estimating the optimal amount of such devices and their respective positions, in order to minimize the FL errors.

The field equipment data and IT systems information follow different protocols, according to their respective manufacturers. Locating faults based on those resources demands proper data integration, which should be addressed by future works aiming at overcoming such specific obstacles.

Based on this research, the authors are involved in an actual FL system implementation, considering some field equipment alarms and readouts as the FL inputs. It will be possible to validate the proposed FL methodology and identify practical challenges regarding the equipment recordings accuracy, availability and data acquisition.

## 5 Appendix

### 5.1 Evolutionary Algorithms Principles

The evolutionary algorithms are inspired by biological evolutionary theory. According to that theory, the individuals in nature interact with each other and with their environment. The fittest ones to the environment are more likely to survive, reproduce and then generate descendants (Von Zuben 2000).

The main idea of evolutionary computation algorithms is to regard the solution of a complex problem as the evolution process presented by sets of individuals, in which each individual represents a particular solution for the problem.

```
t = 0
initialize P(t)
evaluate P(t)
While (Stop criterion not satisfied) do
    P'(t) = variation P(t)
    evaluate P'(t)
    Q(t) = f[P(t)]
    P(t+1) = selection [P'(t) U Q(t)]
    t = t + 1
End
```

**Fig. 7** Evolutionary strategy steps

The steps for the individuals evolution, which are conducted by computational implementation, are show by Fig. 7. According to the steps description, a population $P(t)$ with individuals of generation $t$ suffers variations, generating population $P'$, whose individuals are then evaluated. A subset $Q(t)$ of individuals from $P(t)$ and those from $P'$ are selected for the following generation. The whole process is repeated, until the stop criterion is satisfied.

In evolutionary strategies (ES), each individual is characterized by a vector of parameters $\mathbf{x} = (x_1, \ldots, x_n)$, which varies according to the standard deviations vector $\boldsymbol{\sigma} = (\sigma_{x_1}, \ldots, \sigma_{x_n})$, where $n$ is the amount of parameters. Within a population, the individuals are affected by the ES operators, which will be described later. An individual $\mathbf{x} = (x_1, \ldots, x_n)$ is evaluated through an evaluation function $f_{\text{eval}}(x_1, \ldots, x_n)$.

### 5.2 Operators

From an individual characterized by $(\mathbf{x}, \boldsymbol{\sigma})$, mutation operator generates descendants whose parameters $(\mathbf{x'}, \boldsymbol{\sigma'})$ are calculated according to Eqs. 9 and 10.

$$\sigma_i' = \sigma_i \cdot \exp(\tau' \cdot N(0,1) + \tau \cdot N_i(0,1)) \tag{9}$$

$$x_i' = x_i + \sigma_i' \cdot N_i(0,1) \tag{10}$$

where

$\sigma_i$: parameter $x_i$ mutation step
$\sigma_i'$: new value of $\sigma_i$
$N_i(0,1)$: random from normal distribution with $(\mu = 0, \sigma = 1)$
$N(0,1)$: same of $N_i(0,1)$, but constant for the individual
$\tau_i'$: learning rate of $x_i$. Usually, $\tau_i' = \frac{1}{\sqrt{2\beta}}$, with $\beta = 2$
$\tau_i$: learning rate of $x_i$. Usually, $\tau_i = \frac{1}{\sqrt{2\sqrt{\beta}}}$, with $\beta = 2$

The Crossover operator generates a single descendant from individuals $A$ $(\mathbf{x_A}, \boldsymbol{\sigma_A})$ and $B$ $(\mathbf{x_B}, \boldsymbol{\sigma_B})$, generating a single descendant with parameters $(\mathbf{x_C}, \boldsymbol{\sigma_C})$, where $a$ is a random number in the range $(0, 1)$ and $i$ is the parameter index (Kagan et al. 2009).

$$x_{C_i} = a \cdot x_{A_i} + (1 - a) \cdot x_{B_i} \tag{11}$$

$$\sigma_{C_i} = a \cdot \sigma_{A_i} + (1 - a) \cdot \sigma_{B_i} \tag{12}$$

The Selection operator defines which individuals will move to the following generation. Firstly, all the individuals are ordered in terms of value obtained by the evaluation function. The first $N_{\max}$ individuals are selected for the following generation, where $N_{\max}$ is the maximum number of individuals in a generation.

## References


Al-shaher, M. A., Sabry, M. M., & Saleh, A. S. (2003). Fault location in multi-ring distribution network using artificial neural network. *Electric Power Systems Research*, *64*(2), 87–92.

Cordova, J., & Faruque, M. O. (2015). Fault location identification in smart distribution networks with distributed generation. In *North American power symposium (NAPS)*, 2015, IEEE (pp. 1–7).

Džafić, I., Jabr, R. A., Henselmeyer, S., & Donlagić, T. (2018). Fault location in distribution networks through graph marking. *IEEE Transactions on Smart Grid*, *9*(2), 1345–1353.

Farias, P. E., de Morais, A. P., Junior, G. C., & Rossini, J. P. (2016). Fault location in distribution systems: A method considering the parameter estimation using a rna online. *IEEE Latin America Transactions*, *14*(12), 4741–4749.

Girgis, A. A., & Fallon, C. M. (1992). Fault location techniques for radial and loop transmission systems using digital fault recorded data. *IEEE Transactions on power delivery*, *7*(4), 1936–1945.

Guerra, W., & Kagan, N. (2009). Fault locations in transmission systems by evolutionary algorithms. In *ICREPQ*.

Kagan, N., Schmidt, H., Oliveira, C., & Kagan, H. (2009). *Métodos de otimização aplicados a sistemas elétricos de potência*. São Paulo: Blucher.

Manassero, G., Di Santo, S. G., & Souto, L. (2017). Heuristic method for fault location in distribution feeders with the presence of distributed generation. *IEEE Transactions on Smart Grid*, *8*(6), 2849–2858.

Parker, D., & McCollough, N. (2011). Medium-voltage sensors for the smart grid: Lessons learned. In *Power and Energy Society General Meeting*, 2011 IEEE, IEEE (pp. 1–7).

Saha, M. M., Das, R., Verho, P., & Novosel, D. (2002). Review of fault location techniques for distribution systems. *Power systems and communications infrastructures for the future, Beijing*.

Senger, E. C., Manassero, G., Goldemberg, C., & Pellini, E. L. (2005). Automated fault location system for primary distribution networks. *IEEE Transactions on Power Delivery*, *20*(2), 1332–1340.

Trindade, F. C., Freitas, W., & Vieira, J. C. (2014). Fault location in distribution systems based on smart feeder meters. *IEEE Transactions on Power Delivery*, *29*(1), 251–260.

Von Zuben, F. J. (2000). Computação evolutiva: uma abordagem pragmática. Tutorial: Notas de Aula da disciplina IA707, Faculdade de Engenharia Elétrica e de Computação-Universidade Estadual de Campinas.

Zhu, J., Lubkeman, D. L., & Girgis, A. A. (1997). Automated fault location and diagnosis on electric power distribution feeders. *IEEE Transactions on Power Delivery*, *12*(2), 801–809.