FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Neural Networks Improving the Performance of the Distance Protection

Luís Miguel Andrade Barreira

Dissertation integrated in the Master Degree in Electrical and Computer Engineering
Major: Energy

Supervisor: Prof. Dr. Vladimiro Henrique Barrosa Pinto de Miranda
Full Professor at FEUP

Co-Supervisor: Prof. Dr. Hélder Filipe Duarte Leite
Assistant Professor at FEUP

June 2013

Abstract

The distance protection, or distance relay, is a type of protection system responsible for detecting faults in power system lines and adopting the necessary actions to isolate the fault. This device contributes to the security and reliability of the system, avoiding loss of load or loss of synchronism in case of disturbances. Therefore its correct operation is of major importance.

The operation of distance relays is influenced by different internal and external factors, compromising its performance. To improve the accuracy of these systems new solutions are being studied.

Then, in this dissertation the influence of external factors; namely pre-fault load condition, intermediate in-feeds and heavy load; is studied and a neural network based fault location scheme is proposed to improve the performance of the distance protection. To create and evaluate the performance of the solution proposed a data set is generated encompassing different pre-fault and fault conditions for a given test system. The data set created is then used for training the neural networks. To finalize the performance of the solution proposed is compared with the performance of a mho relay.

The results show that neural networks are an efficient tool for improving the distance protection performance.

Key Words: Distance Protection, Pre-Fault Load Flow, In-feed, Heavy Load, Neural Network

Acknowledgements

First of all I want to thank professor Vladimiro Miranda for the opportunity to work with him, for his guidance, commitment and endless ideas, but most of all, for his belief in my work.

My sincere thanks also goes to professor Hélder Leite for being there any time that I needed and for his efforts in helping me understand and write this dissertation.

I also want to thank Joana da Hora who supported me all along, enlightening me in the world of neural networks.

A very kind word goes to my parents and brother, for their support and care, that helped me accomplish this stage and, in particular, this work.

To all my sincere gratitude,

Luís Miguel Andrade Barreira

Contents

List of Figures

List of Tables

Abbreviations and Symbols

List of abbreviations (ordered alphabetically):

ANFIS	Adaptive Neurofuzzy Inference System
FNN	Feedforward Neural Network
NN	Neural Network
PCA	Principal Component Analysis
RCA	Relay Characteristic Angle
SCADA	System Control and Data Acquisition

Chapter 1

Introduction

1.1 Problem Specification

Power systems are subjected to several disturbances during normal operation. The correct detection of these disturbances and adoption of the necessary recovery actions is of major importance, so that the system regains its stable operating point without occurring loss of load or, in the worst case scenario, loss of stability. This function is performed by protective devices.

The distance protection, or distance relay, is one of these devices. Its main application is in the protection of transmission lines, therefore the objective of the distance protection is to detect line faults and isolate, as quickly as possible, the faulted line. The correct fulfilment of this objective depends on a number of factors related to the measuring accuracy of the elements that compose the distance protection and its setting by the protection engineer.

The relay operates incorrectly when the measured impedance is different from the real short-circuit impedance between the relay location and the fault location or when the impedance seen by the relay during normal operating condition is less than a certain setted threshold value. The incorrect operation of the relay may be classified into two different situations[7]:

- **Underreach**: the relay does not operate for a disturbance inside its protection zone. The relay does not operate when it should.
- **Overreach**: the relay operates for disturbances external to its protection zone. The relay operates when it should not.

Different causes may be in the origin of underreach and overreach situations. Among these, the influence of intermediate in-feeds, pre-fault load conditions and heavy load conditions are the problems in study in this dissertation:

- **Intermediate In-feeds**
  The intermediate in-feeds consist of injections of current, located between the relay

location and the fault location, that influence the apparent impedance measured by the relay leading the relay to operate incorrectly.

- **Pre-fault Condition**
  As the intermediate in-feeds, the pre-fault load condition may cause a deformation in the impedance measured by the relay, provoking an incorrect operation.
- **Heavy Load Condition**
  When the system is under an heavy load condition the relay may operate incorrectly since the apparent impedance measured may be under a setted threshold value that defines the protective reach of the relay.

1.2 Methodology

The final objective of this dissertation is to propose a solution, based on neural networks, capable of improving the performance of distance protection. To do so, one must understand the distance protection, know its operating errors, analyse this errors and, finally, create an adequate solution. The development process of this dissertation is as follows:

1. **Overview on Distance Protection** (see chapter 2)
   The first phase is basically understanding all aspects related to the distance protection, namely the importance of its correct operation, its principles of operation and operating errors. Besides a review on the applications developed to improve the operation of the distance protection is also presented.
2. **Distance Protection Simulated Study** (see chapter 3)
   Once the problem is specified and the works developed on distance protection are reviewed, the next step is to analyse the behaviour of a distance relay installed on a given test system. To do so a simulation is created, whose objective is to sample different pre-fault conditions and fault locations on the test system (in this simulation only three phase faults were considered). With the analysis of the test system and simulation results the problem in study is characterized. Besides the incorrect operations of the relay are identified for a certain data sample generated with the simulation created.
3. **Improving the Performance of Distance Protection** (see chapter 4)
   In the final phase the solution developed is presented. This solution is modelled based on the results of the analysis carried out and on the incorrect relay operations identified in the data set generated. Since the operation of the relay is basically a classification procedure, and it is known that neural networks have the capacity of classifying patterns in a more precise way [8], this method is the basis of the solution proposed. Therefore the adequacy of neural networks to discriminate non fault from fault conditions is evaluated. Besides their capacity of establishing a relation between the pre-fault and fault conditions is tested. In short, this phase involves presenting the building blocks of the

neural network based distance protection, analysing its capacity for solving the problems in study and, finally, determining if this solution improves the relay operation.

1.3 Research Questions and Thesis

- What is the influence of the pre-fault condition on the impedance measured by the distance protection?
- What is the influence of in-feeds in the impedance measured by the distance protection?
- Which are the errors in the distance protection performance due to in-feeds, pre-fault conditions and heavy load conditions?
- Can neural networks learn the relationship between pre-fault and short-circuit conditions?
- Can neural networks improve the performance of the distance protection?

Thesis: computational intelligence techniques may improve the performance of the distance relay.

Chapter 2

Overview on Distance Protection

2.1 Power System Protection

Electrical Energy is one of the main power resources of the contemporary society. Electrical power availability is a critical necessity: the power system must be able to supply the exact amount of energy needed at the correct voltage and frequency. Frequent or prolonged interruptions in power supply can cause severe disruptions in modern social routine. To achieve the desired standards of economy, reliability and security a careful planning, design, installation and operation is needed. With the growing installed capacity and power system complexity the requirements of reliability and economy present an even greater challenge and make the power system design a compromise [4].

To the consumer of electricity the power appears to be always available, yet the operation of power systems is constantly being subjected to disturbances derived from load changes, faults created by natural causes, equipment damage or operator failure. In most disturbances cases the power system is able to maintain its quasi-steady state due to its large dimension (when compared to the size of individual loads and generators) and correct operation of protection devices [7] .

Power system protection is the area of power engineering concerned with the design, implementation and operation of protection devices, called 'relays' or 'protective relays'. These devices have the function of detecting abnormal power system conditions and adopt the necessary actions, as rapidly as possible, in order to return the power system to its normal operation mode [7].

To achieve the desired performance the relays must fulfil the following requirements [4]:

- **Sensitivity**: this characteristic is related to the relay minimum operating level, the lower this parameter is the more sensible the relay becomes.
- **Speed**: is the relay capacity to isolate faults as quickly as possible, minimizing the damage in power system equipment, safeguarding the continuity of supply and avoiding the loss of synchronism and consequent collapse of the power system.

- **Stability**: is the ability of the relay to remain unaffected by occurrences outside its protection zone, namely external faults or heavy load conditions.
- **Selectivity**: is the ability of isolating only the faulted zone.

The optimal implementation and coordination of protective relays is obtained taking into account the objectives referred above, the topology of the system to be protected, the typical scenarios of operation and the probable fault occurrences.

2.2 Protection of Transmission Lines

Transmission systems are interconnected circuits, composed by meshed lines, with normally more than one source of voltage. Usually, the topology of these systems increases the difficulty of coordinating protective devices since the current may flow in different directions depending on the location of the fault[2]. The protection of transmission lines may be accomplished by different types of relaying techniques, namely the directional overcurrent relay, the pilot relay and the distance relay:

- **Directional Overcurrent Relay**
  Overcurrent relays are the simplest and cheapest form of transmission lines protection, however the correct coordination of these devices is the most difficult to achieve[2]. These type of relays are also very susceptible to the relative source impedances and system condition [1].

- **Pilot Relay**
  Pilot relays operating principle is based on the combination of the observations from both ends of the protected line. This protection concept provides an increase on selectivity, which leads to a greater security of operation. The downside of these systems is that they cost more than the ones where the trip decision is made with local measurements only[2].

- **Distance Relay**
  Distance relays represent the first option for replacing overcurrent relays when these are considered inadequate for a certain function [2]. These type of relaying devices are not affected nearly as much as overcurrent relays to relative source impedances and system conditions[1]. Other advantages offered by distance relays are the integrated fault location function, the possibility of being applied as a remote back-up protection[1] and the wide variety of characteristics, which make the option for these devices the most adequate for certain applications [2].

2.3 Distance Protection

Distance protection is the basis of transmission lines protection. The distance relay is non-unit protection device whose mode of operation is based on the comparison of a measured short-circuit impedance, which is proportional to the distance to the fault, with a pre-defined impedance value [1].

2.3.1 Operating Principle

The impedance of a transmission line is proportional to its length, so by determining the fault impedance, from the measured short circuit voltage and current at the relay location, is possible to identify the location of the fault. This capacity allows the relay to discriminate faults that occur in different lines or line sections[4].

The basic principle behind the impedance measuring is the division of the voltage and current signals, that come from the intensity transformers at the relay location (see figure 2.1). The trip command is issued based on the comparison of the apparent impedance measured by the relay with the known line impedance: if the measured impedance is smaller then the setted line impedance a trip order is issued[1].

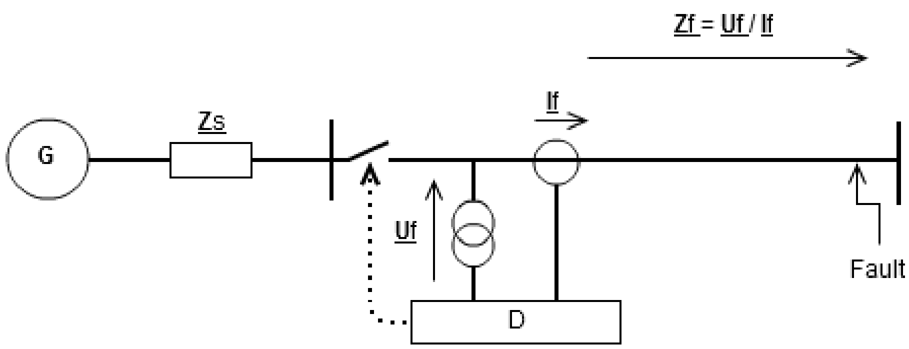

Figure 2.1: Distance protection principle. Adapted from [1].

The impedance measured during normal operation has a magnitude inversely proportional to the amount of transferred load and an angle equal to the ratio between the real and reactive power transferred. When a fault occurs the measured impedance becomes the short-circuit impedance. This value is usually smaller than the load impedance and matches the line impedance between the relay location and the fault location. When the fault has a resistance component, resulting form arc resistance or ground resistance, an additional resistive component is added to the line impedance. The angle corresponds to the lag between the short-circuit voltage and current angles[1].

$$\begin{aligned} \textit{Load impedance magnitude}: \quad |Z_{load}| &= \frac{U_{line}^2}{P_{load}} \quad [\Omega] & (2.1) \\ \textit{Load impedance angle}: \quad \varphi_{load} &= \arctan\frac{P_{reactive}}{P_{real}} \quad [^\circ] & (2.2) \\ \textit{Short-Circuit impedance}: \quad Z_{load} &= Z_{line} + R_{fault} \quad [\Omega] & (2.3) \end{aligned}$$

$U_{line} = voltage\ at\ the\ relay\ location\ [V]$
$P_{load} = \ power\ consumed\ [VA]$
$P_{reactive} = reactive\ power\ consumed\ [var]$
$P_{real} = active\ power\ consumed\ [W]$
$Z_{line} = line\ impedance\ [\Omega]$
$R_{fault} = fault\ resistance\ [\Omega]$

2.3.2 Distance Zones

The impedance measurement has associated inaccuracies resulting from intensity transformers and relay transformation errors, line impedance calculation, line impedance variations derived from atmospheric conditions, source impedance variations and pre-fault operating condition, so it is not possible to set the protection reach to 100% of the line[1]. On the other hand the capacity of detecting faults on lines protected by other equipments offers the possibility of back-up protection. The combination of these factors is achieved by introducing a stepped distance protection concept which is basically dimensioning different zones with an associated reach and tripping time (see figure 2.2). The definition of different zones allows a discriminated tripping and enables the correct coordination between distance relays on the power system.

Modern microprocessor relays available in the market have up to 5 zones of protection [9] [10] [3]. The common description of each zone is [5]:

- **Zone 1**: Instantaneous tripping zone that is normally set to provide a forward reach up to 80 - 90% of the protected line.
- **Zone 2**: Time delayed tripping zone that covers the protected line and provides back-up protection to the next line. This zone has the objective of detecting faults in the protected line beyond Zone 1 and offering remote back-up protection for a failed zone 1 element both in the protected line and next line. Zone 2 reach is normally setted to be 120% of the protected line. A higher setting can be defined if the fault in-feed from adjacent lines at the remote end is considerably higher than the fault current at the relay location. This setting shall not be higher than 80% of the impedance of the protected line plus the first zone reach of the shortest adjacent line. If the protected line

zone 2 element reach cannot be setted to less than the zone 1 reach of the next line, a coordination is achieved by an adjustment on the tripping time of one or both zone 2 elements.

- **Zone 3**: Forward-looking zone with the purpose of offering remote back-up protection to clear a fault if a remote breaker does not trip. This is a time delayed zone setted to offer protection to 100% of the largest adjacent line in maximum in-feed conditions. Sometimes zone 3 setting becomes high enough to operate on load condition, voltage instability or power swings. To avoid the undesired operation of the relay in such cases other measures are adopted which encompass using shaped characteristics, load encroachment detection and power swing blocking.

- **Reverse Zone 3 and off-set zones**: These are time delayed zones that may be used to offer reverse back-up protection to elements behind the relay location or an offset forward-looking protection zone. The purpose of these zones is providing a scheme communication logic, current reversal logic, weak-end in-feed logic, and so on.

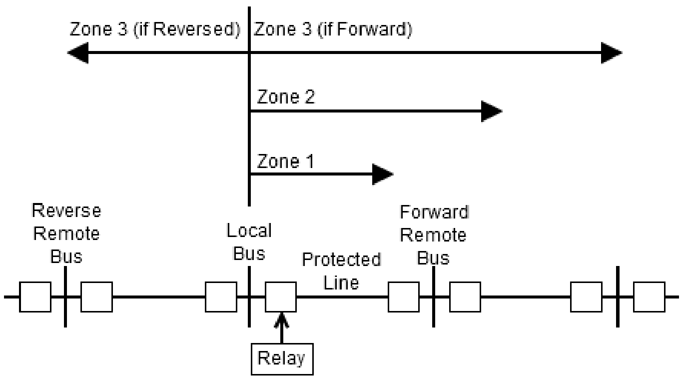

Figure 2.2: Zones reach. Adapted from [2].

2.3.3 Impedance Diagram

The impedance diagram is the primary tool to setting and evaluating the relay behaviour. Since the impedance measured by the relay is the ratio of voltage and current and the angle between them, it represents a point in the complex R-X plane. In this diagram, the relay operating characteristic for the different tripping zones is represented. By studying the relation between the loci of power system impedances during normal operation, faults and disturbances the protection engineer can determine the adequate setting of the zones[4].

The operating characteristic consists of a fixed shape in the R-X plane which defines the separation between the load and fault areas. The traditional relay operating characteristic shapes are geometric figures composed by straight lines and circles or sectors of circles (see figure 2.3).

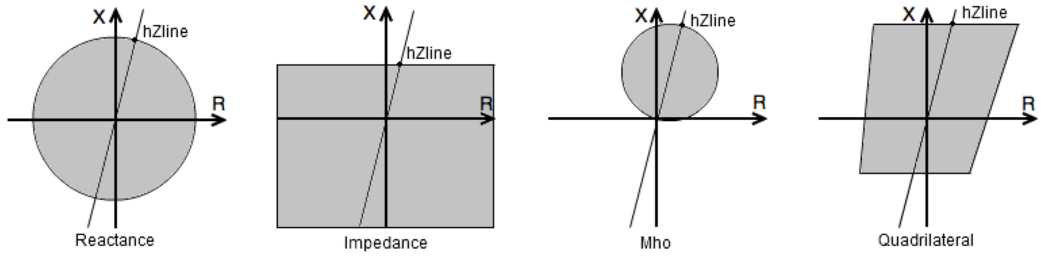

Figure 2.3: Operating characteristics in the R-X plane. Adapted from [1].

The trip decision discrimination between different areas of the system is obtained by defining different zones with different reaches.

Another important factor is the directional measurement. In presence of multi-terminal radial systems or meshed systems it must be determined whether the fault is in the forward or reverse direction to avoid incorrect tripping for reverse faults that occur outside the protected line. The determination of the fault direction can be shown in the R-X plane: for faults occurring in the forward direction the current flows forward and the angle is usually above 80° for EHV overhead lines and 20° for cables; for faults occurring in the reverse direction the current is reversed and the angle appears to be rotated 180° when compared to forward direction faults[1] (see figure 2.4).

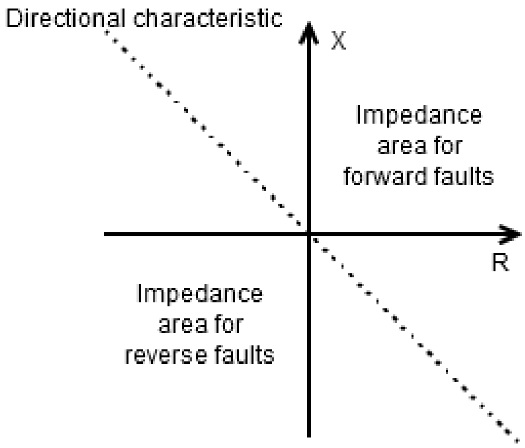

Figure 2.4: Directional characteristic in the R-X plane. Adapted from [1].

Nowadays, with the increase of processing power, relays have available different operating characteristics and allow their optimization. The most common characteristics in commercial relays are mho and quadrilateral [9] [10] [3] (see figure 2.5).

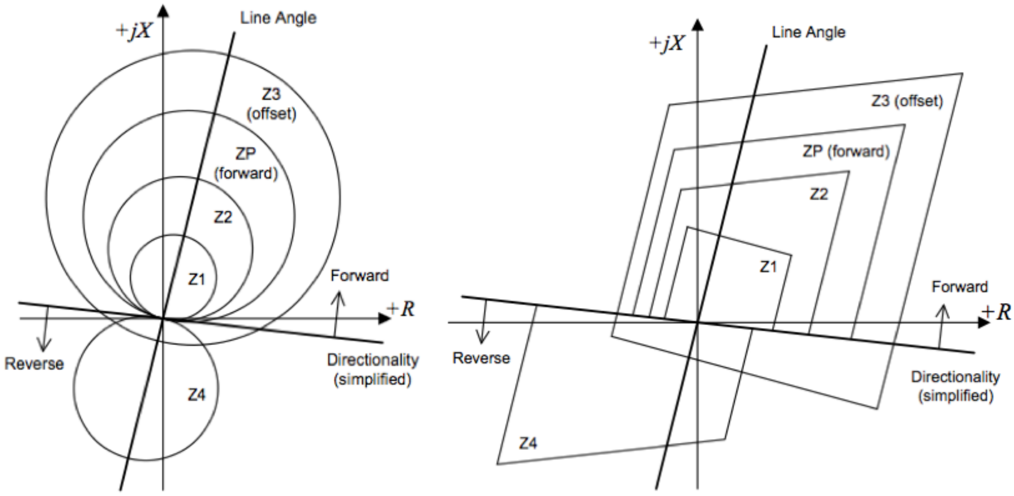

Figure 2.5: Left: mho characteristic; right: quadrilateral characteristic[3].

The mho characteristic, when plotted in a R-X plane, consists of a circle whose circumference passes through the origin, this demonstrates that the relay offers both directional and reach control. The adjustment of this characteristic is made by setting the impedance reach along the diameter of the circle (AQ) equal to the intended line impedance coverage and the angle of displacement of the diameter from the R axis (φ) equal to the line impedance angle (see figure 2.6). The angle φ is known as the relay characteristic angle (RCA) or the torque angle. The relay operates when the fault impedance is inside the circumference.

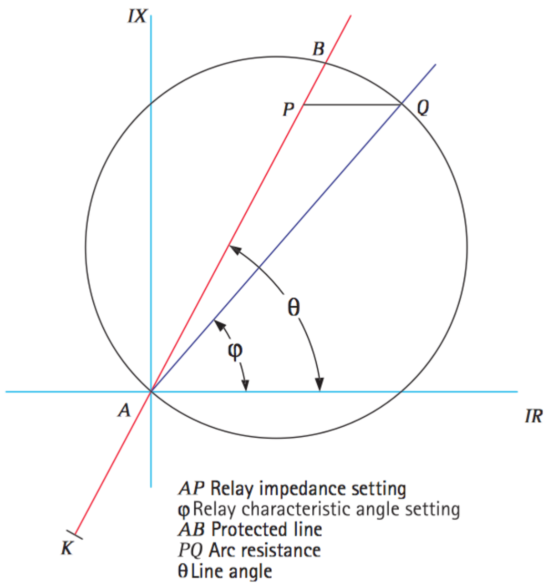

Figure 2.6: Mho characteristic settings[4].

The fault angle depends on the relative values of R and X of the line impedance plus the arc fault or earth fault resistance. Thus, a relay having a characteristic angle equivalent to the line angle may not trip when facing high resistive faults because the resistive component will move the impedance point to the region outside the protection zone. To avoid incorrect trip decisions the RCA is usually set to less than the line angle. The setting of this characteristic is shown in figure 2.6, where AB is the impedance magnitude of the line to be protected and AQ is the diameter of the circle which is obtained by dividing the actual amount of protected line by the cosine of the subtraction of the relay characteristic angle θ by the line angle φ (to adopt this setting the difference between this to angles must be known)[4]. Hence, the relay setting AQ is given by:

$$AQ = \frac{AB}{cos(\theta - \varphi)} \tag{2.4}$$

The quadrilateral characteristic has an independent adjustment of forward reach and resistive reach which gives a better resistive coverage than any mho type and makes them a specially adequate to prevent the relay incorrect operation in the presence of arc resistance faults and earth faults. The setting of this characteristic is shown in figure 2.7 where AB is the impedance magnitude of the line to be protected and BC is the impedance magnitude of the longest adjacent line.

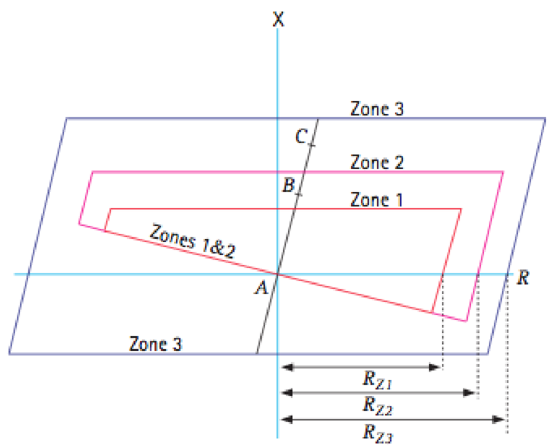

Figure 2.7: Quadrilateral characteristic settings[4].

2.4 Distance Protection Incorrect Operations

Traditionally, all relay settings are a compromise. This characteristic makes the distance protection relay susceptible to incorrect operation situations, since the impedance measurement has an error associated. These situations can be divided in two groups[7]:

- **Underreach**: the relay does not operate for a disturbance inside its protection zone. The relay does not operate when it should.
- **Overreach**: the relay operates for disturbances external to its protection zone. The relay operates when it should not.

Underreach and overreach events may lead to equipment damage, power system loss of stability or cascaded blackout events. The origin of these situations is associated with the equipment measuring error and external system conditions [1]:

- Precision of the relay and intensity transformers.
- Saturation of intensity transformers.
- Transitory state measuring error of VT's.
- Inaccuracies in the line impedance calculation.
- Line impedance variations derived from atmospheric conditions.
- Pre-fault load condition.
- Intermediate In-feeds.
- Heavy load.
- Source impedance ratio.
- Fault resistance.
- Mutual coupling effect.
- Series compensated lines.

In this thesis will only be addressed the undesired operations due to pre-fault load condition, in-feeds and heavy load.

2.4.1 Pre-fault Load Condition

In [11] the authors have studied the influence of initial power flow condition and system topology influence in the impedance measuring. They have concluded that the pre-fault active and reactive power flow is one of the main influencing factors in the measurement of the fault impedance. On the other hand, the system topologies studied, based on the variation of lines length, have not presented significant impact.

The solution proposed consists in the dimensioning of an ideal operating region that takes into account the normal range of active and reactive power flow based on the usual system configuration. For future study is proposed the development of an intelligent relay that is able to adapt its operating region to system pre-fault active and reactive power flow.

2.4.2 In-feed Effect on Different System Topologies

In-feed consists of an injection of current. Figure 2.8 contains three system topologies where the in-feed condition presents different effects (see figure 2.9). The first topology (a) consists of a radial system with an intermediate in-feed, the second (b) represents a parallel line system and the third (c) represents a simplified meshed system[1].

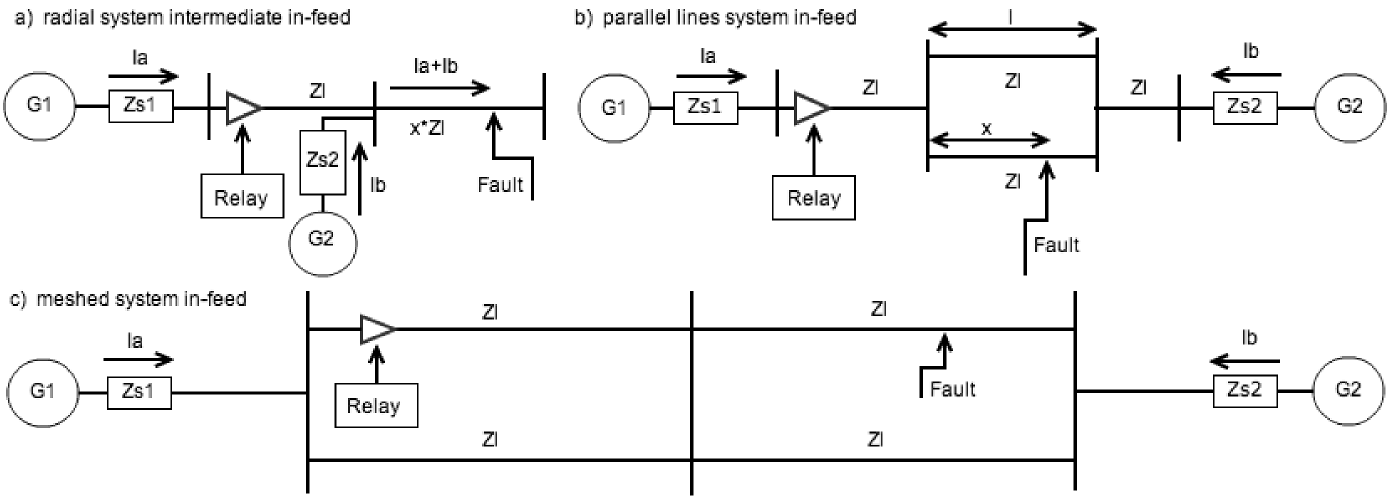

Figure 2.8: In-feed in different system topologies. Adapted from[1].

In case a), where the in-feed is located between the relay location and the fault location, the impedance seen by the relay will be greater than the real fault impedance and the relay will underreach. This phenomena occurs because the intermediate current injection introduces an additional voltage drop in the short-circuit loop. The error inherent to the intermediate in-feed condition is proportional to the ratio between the current injected and the current at the relay location: the larger the in-feed, the greater the error. The impedance variation seen by the relay in the case of a radial system with intermediate in-feed is linear and proportional in the R and X axis[1]:

$$Z \quad = \quad Z_l + Z_l * x * (1 + \frac{I_b}{I_a}) \tag{2.5}$$

In case of parallel line in-feed (b) the impedance variation is non-linear, it consists of a parabolic trace resulting from the parallel connection of the impedances $\frac{x}{l}.Z_2$ and $Z_3 + (1 - \frac{x}{l}).Z_2$.

In meshed systems the described effects of parallel line topologies and intermediate in-feeds are combined.

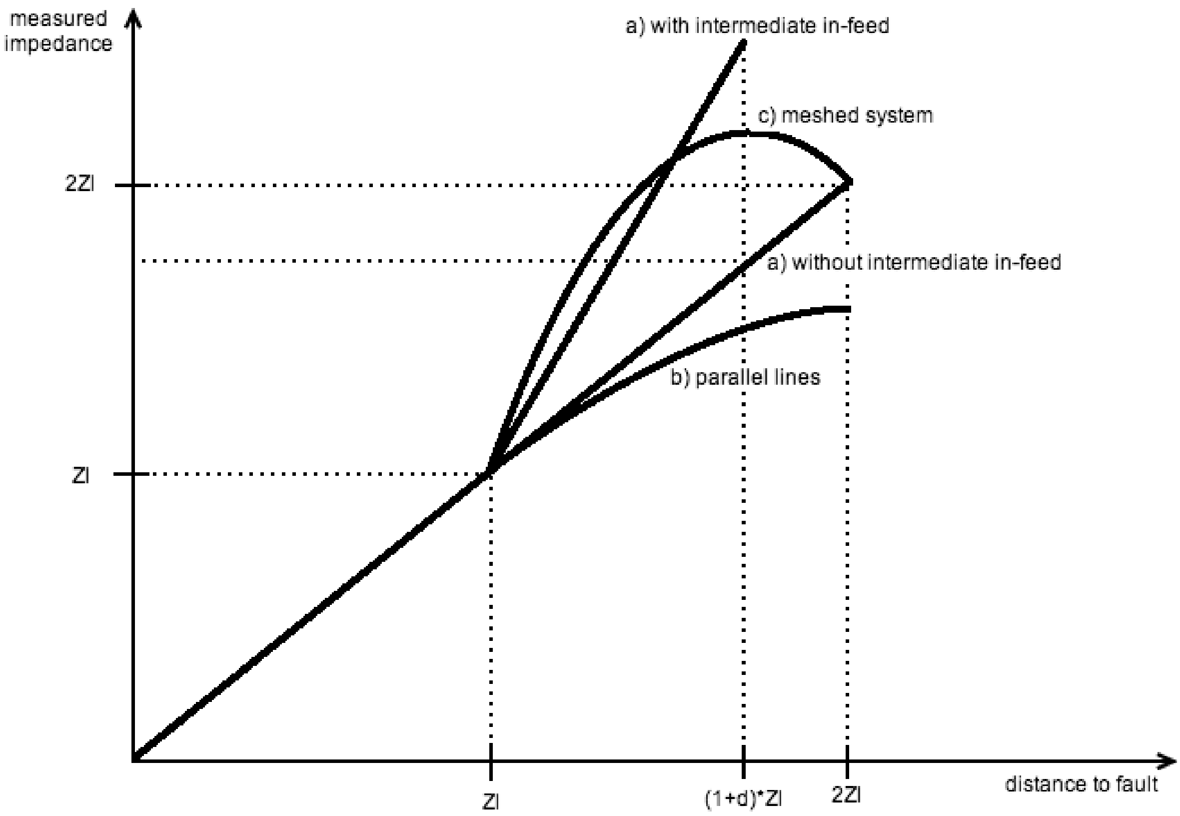

Figure 2.9: In-feed effects on different system topologies. Adapted from [1].

The identification and calculation of in-feeds is of critical importance. To correctly set the protection zones of the relay these effects must be taken into account[1][7]:

- **Zone 1 setting**: in zone 1 the in-feed effect is only present on T-eed feeders. In this case the reach of the zone must be increased to take into consideration the effect of the injected current in the impedance measuring.

- **Zone 2 setting**: in zone 2 the in-feed effect is not taken into consideration.

- **Zone 3 setting**: in zone 3 there are two alternatives: intermediate in-feeds are taken into account or not taken into account. The first option results in large reaches in zone 3, which may lead to the unintended operation of the relay during heavy load conditions, in the case of several in-feeds the approach of considering only part of them may be more accurate. The second option presents very short 3^{rd} zones, in this case the protection or breaker operation failures are cleared by fault detectors. This approach presents longer fault clearing times but avoids zone 3 unintended operations.

2.4.3 Relay Loadability

As it has been stated in subsection 2.3.3 the impedance seen by the relay during normal operation is related to the load impedance, that during normal operation is far from the relay operating zone (the typical operating points have higher resistive and lower reactive reaches). As the load on the system increases, the apparent impedance locus in the R-X diagram approaches the operating zone of the relay. The relay loadability is the value of maximum load at which the relay is on the verge of operation. For some system heavy load operating conditions, line loadings may occur where the apparent impedance locus crosses the limits of the relay operating zone and the relay trips[7]. Other phenomena like voltage instability or power swing may also lead to an improper zone 3 operation, but these phenomena are out of this dissertation scope.

As the reach increases, the relay trip on heavy load conditions becomes a more probable situation. The factors that influence the relay loadability are:

- The power factor of the line current: more inductive currents move the apparent impedance locus to an higher point in the reactance axis, typically entering zone 3.
- In-feed: this effect leads the relay engineers to adopt longer reaches in zone 3, making the relay more susceptible to heavy load tripping conditions.

To avoid the incorrect operation of the relay due to heavy load the traditionally used methods consist in the alteration of the zone 3 characteristics or in the use of load encroachment zone [5]:

1. Increase of the torque angle.
2. Adoption of a rectangular characteristic.
3. Adoption of a lens characteristic.
4. Use of blinders.
5. Use of load encroachment.

The options 1, 2, 3 and 4 reduce the susceptibility of the relay response to heavy load conditions but diminishes the coverage of resistive faults. Option 5 presents the most effective and reliable method of discriminating system faults from heavy load conditions. The load encroachment is a protection feature available in modern line relay packages (for example:[9] [10] [3]) that enables the protection engineer to define custom load regions in forward and reverse direction. Besides its operation has a supervision restriction that only blocks the relay tripping for 3-phase distance elements. Since load is mainly a balanced 3-phase condition the load encroachment will correctly block the relay from tripping on heavy load conditions and enable its operation for unbalanced faults that occur inside the load region. This protection feature operates correctly for every kind of fault except resistive 3-phase faults corresponding to a positive sequence impedance that occur inside the load region. The load encroachment is the most effective method currently in use [5] (see figure 2.10).

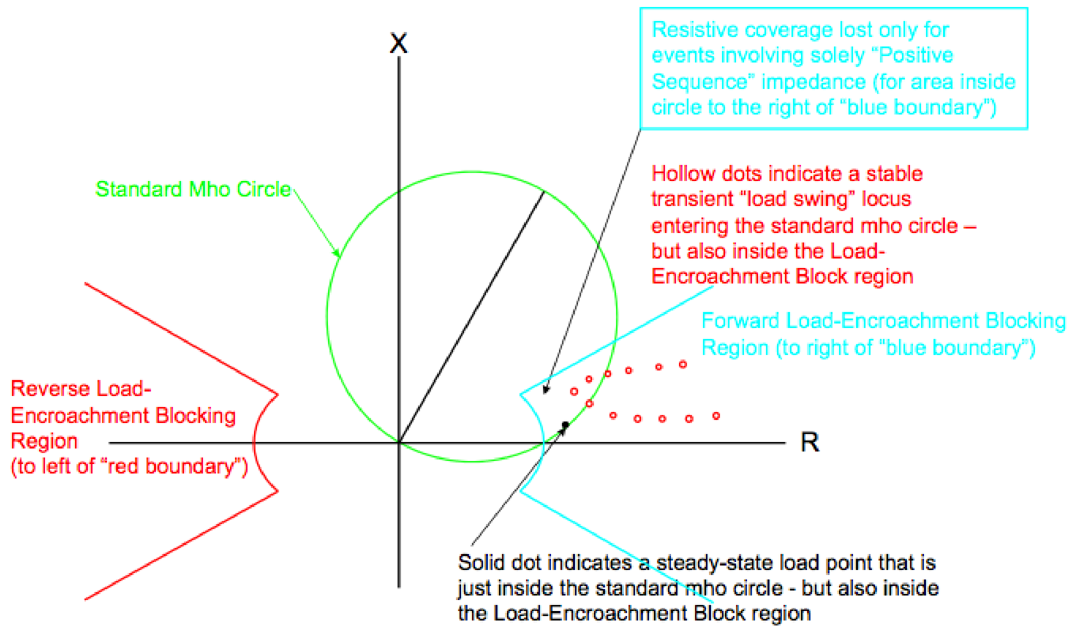

Figure 2.10: Load encroachment feature[5].

2.5 Literature Review on Distance Protection

Distance protection has been an object of study under different perspectives that aim to improve its operation. The description of the different methodologies and approaches adopted in each document are in appendix A. The documents are grouped by different themes, these being:

- **Fault Classification** (A.1)
  The problem of fault classification consists in the identification of the type of fault and phases involved. The accurate and fast fault classification is of major importance since the relay correct impedance measuring depends on the type of fault. This problem has been addressed in [12] , [13] and [14].

- **Fault Distance Estimation** (A.2)
  The problem of fault distance estimation consists in the estimation of the distance between the relay location and the fault. The fault distance estimation is important for the distance relay operation since it is a critical function for fault location. This problem has been addressed in [15] , [16] and [17].

- **Fault Location** (A.3)
  The problem of fault location combines fault classification and location and is studied in [18] , [19], [20] and [21].

- **Adaptive Zone** (A.4)
  The adaptive zone concept defends that by modifying the protection zone, adapting it to the system operating conditions, the number of overreach and underreach situations may be reduced. This approach is studied in [22] , [23], [24] and [25].

- **Zone 3 Unintended Tripping** (A.5)
  The zone 3 unintended tripping consist in the incorrect zone 3 tripping due to heavy load(described in section 2.4.3), voltage instability and power swing. Methodologies to solve this problem are proposed in [26] , [27], [28] and [29].

- **Other Applications** (A.6)
  Besides the problems and methodologies described other approaches have been studied. The documents in [30] and [31] propose the inclusion of the wavelet transform to optimize distance protection operation while in [32] a neurofuzzy inference system (ANFIS) is proposed to detect fault situations.

Chapter 3

Distance Protection Simulated Study

3.1 System Analytical Model

The performance of the relay is conditioned by external and internal operational characteristics. The focus of this dissertation is to analyse the effects of external conditions; namely, the pre-fault condition, intermediate in-feeds and heavy load; on the relay performance and propose an optimized distance protection scheme. For this analysis a data set was created based on the following assumptions:

- The data sample must cover a wide range of possible pre-fault loading conditions and line fault locations.
- The system topology must include parallel lines and intermediate in-feeds.

To obtain a data set respecting all the factors indicated above a simulation was created, inspired on the monte carlo method. This simulation consists of randomly sampling different operating conditions and line faults of a given test system. The test system is a meshed system, composed by three generators, three lines, three power transformers and three loads (the test system characteristics are described in appendix B). The relay dimensioned has an mho characteristic and is located in line 1 (the relay setting is described in appendix C). Its important to refer that only 3-phase line faults were included in this study due to time limitations. The simulation algorithm was implemented in Matlab.

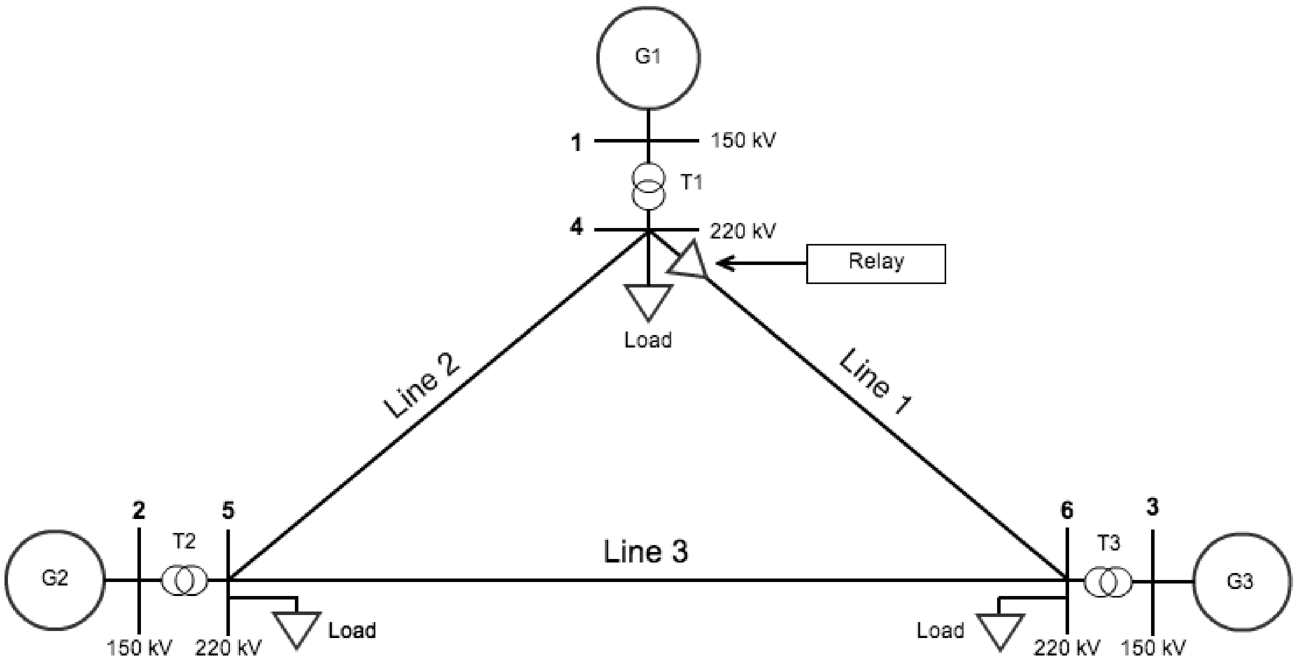

Figure 3.1: Test System.

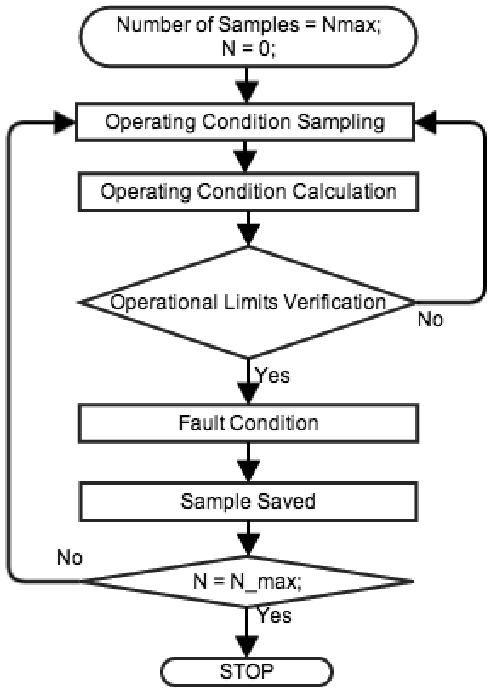

Figure 3.2: Simulation Algorithm.

The algorithm flowchart is presented in figure 3.2 and consists in the following steps:

1. **Operating Condition Sampling**
   This step consist of determining a system operating condition by randomly defining the loading level, the power consumed by each load, the active power generated by each generator and the specified voltage in each source:

   - Loading level: the total active and reactive power consumed by the loads is defined by randomly sampling a number between the system maximum and minimum generating capacity.
   - Load allocation: for each load a randomly sampled percentage of the consumed active and reactive power is assigned. The total power allocated is equal to the system loading level.
   - Generation allocation: for each source, the active power produced is randomly defined taking into account that this value must be within the maximum and minimum limits of the generator and the total active power generated must be equal to the total active power consumed.
   - Specified voltage: for each generator a specified voltage is sampled. This value varies between 0.9 and 1.1 pu.

2. **Operating Condition Calculation**
   In this step the operating condition sampled in 1 is calculated using the Newton-Raphson method and the impedance seen by the distance relay is determined.
3. **Operational Limits Verification**
   To accept the operating condition sampled as a possible hypotheses the following operational limits must be verified:

   - Generator reactive power production limits (the reactive power generated by each generator is a result of the power flow, so it is not limited in the sampling process).
   - Slack bus generator active power production limits (this value is a result of the power flow and depends on the system losses so it is not limited in the sampling process and may exceed the maximum production limit of the generator).
   - Voltage limits at the load buses (the voltage limits considered for the load buses are 0.85 pu and 1.15 pu).
   - Lines capacity (the power flowing through the lines must be within their capacity limit).
   - Transformer capacity (the power flowing through the transformers must be within their capacity limit).

   If all the operational limits are verified the sampled power flow is accepted, otherwise this power flow is discarded and another one is sampled.

4. **Fault Condition**
   The last step includes sampling the location of the 3-phase fault and determining the impedance seen by the relay. The fault sampling process includes determining the line and distance in which the fault occurs. The distance is a value that defines the percentage of line between the bus with lower number connected to the line in question and the fault location. With this approach faults occurring at different points in lines 1, 2 and 3 are sampled encompassing all the possible relay tripping decisions.

3.2 Measured Impedance Behaviour

In this section an analysis of the impedance measured by the relay is conducted considering the simulation algorithm created. This analysis encompasses the study of the in-feed effect and pre-fault condition influence in the relay measured fault impedance. First the variation of the impedance measurement in function of the fault location for the different lines is presented, then studies are conducted using the simulation algorithm created and to finalize this results are analysed. The objective in this section is to fully characterize the causes of incorrect tripping situations on the system simulated.

3.2.1 Impedance Measurement vs. Fault Location

The impedance measurement variation with fault location presents different behaviours for each line, so the mathematical relationship between this two variables is presented for 3-phase faults in lines 1, 2 and 3. This deduction is made using the superposition theorem, despising the effect of the load admittances and assuming that generators G1, G2 and G3 are current sources generating I_1, I_2 and I_3, respectively.

1. **Faults in line 1**.

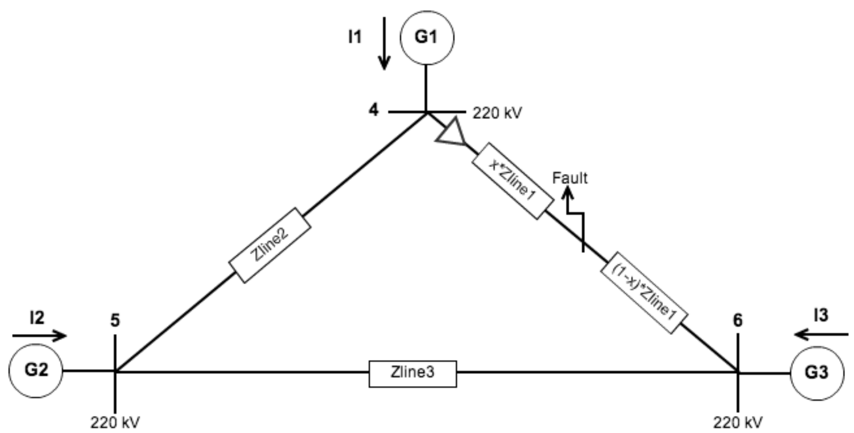

Figure 3.3: Fault Location in line 1.

$$Z_{measured} = Z_{line1} * x \tag{3.1}$$

2. **Faults in line 2**.

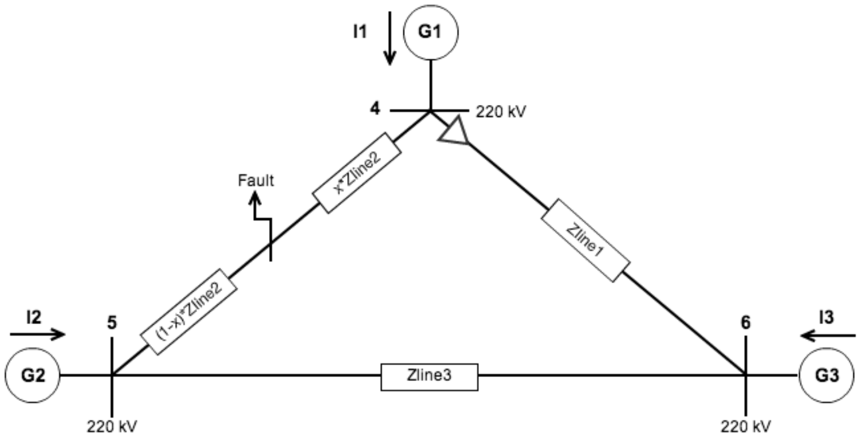

Figure 3.4: Fault Location in line 2.

$$Z_{measured} = Z_{line2} * \frac{-x^2 * Z_{line2} * (I_1 + I_2 + I_3) + x * (Z_{line1} * I_1 + Z_{line2} * (I_1 + I_2 + I_3) + Z_{line3} * (I_1 + I_3))}{x * Z_{line2} * (I_1 + I_2 + I_3) - Z_{line2} * (I_2 + I_3) + Z_{line3} * I_3} \tag{3.2}$$

3. **Faults in line 3**

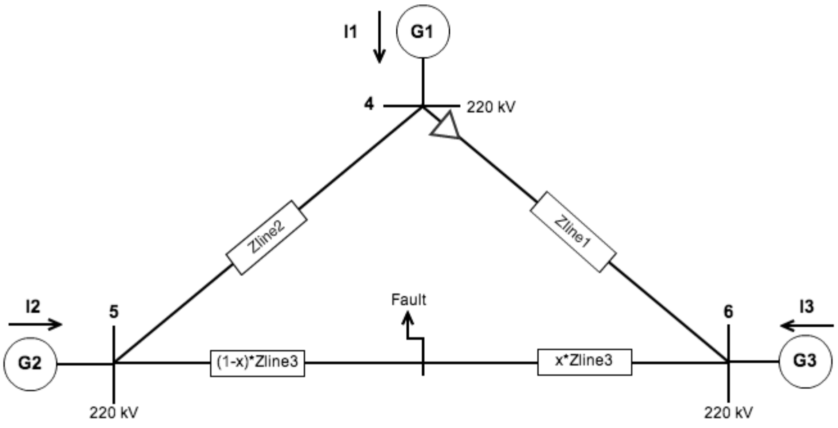

Figure 3.5: Fault Location in line 3.

$$Z_{measured} = Z_{line1} + Z_{line3} * \frac{-x^2 * Z_{line3} * (I_1 + I_2 + I_3) + x * (Z_{line1} * I_3 + Z_{line2} * (I_1 + I_3) + Z_{line3} * (I_1 + I_2 + I_3))}{-x * Z_{line3} * (I_1 + I_2 + I_3) + Z_{line2} * I_1 - Z_{line3} * (I_1 + I_2)} \tag{3.3}$$

3.2.2 Pre-fault and In-feed Conditions Effect

In-feeds in different locations present different behaviours. To study this effect two loading conditions are used to sample stepped distance faults in the whole line length for each line (the step considered is 1%). The loading conditions selected include an high and a low loading levels:

Table 3.1: Load Conditions.

Load Condition	Pgen (MW)			Pload (MW)			Qload (Mvar)			Voltage (kV)		
	Bus 1	Bus 2	Bus 3	Bus 4	Bus 5	Bus 6	Bus 4	Bus 5	Bus 6	Bus 1	Bus 2	Bus 3
High Load	92.6	92.0	120	92.8	142.6	67.8	43.5	66.9	31.8	207.2	203.3	214.0
Low Load	34.2	32.8	32.2	14.2	0.9	82.8	0.8	0.1	4.5	229.4	227.0	240.0

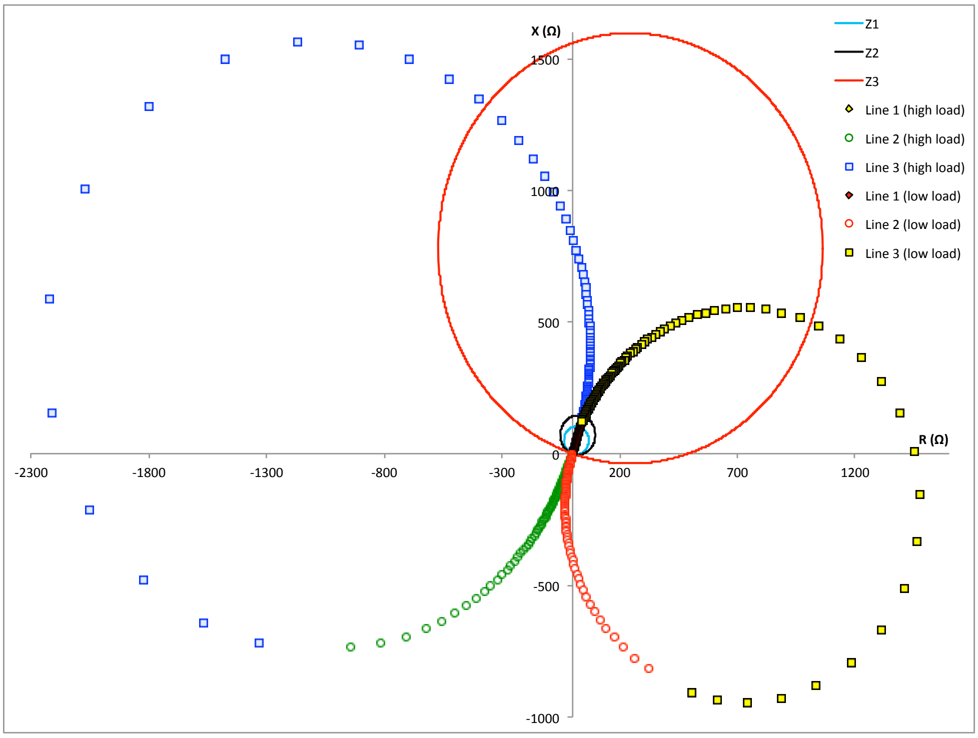

Figure 3.6: In-feed effect.

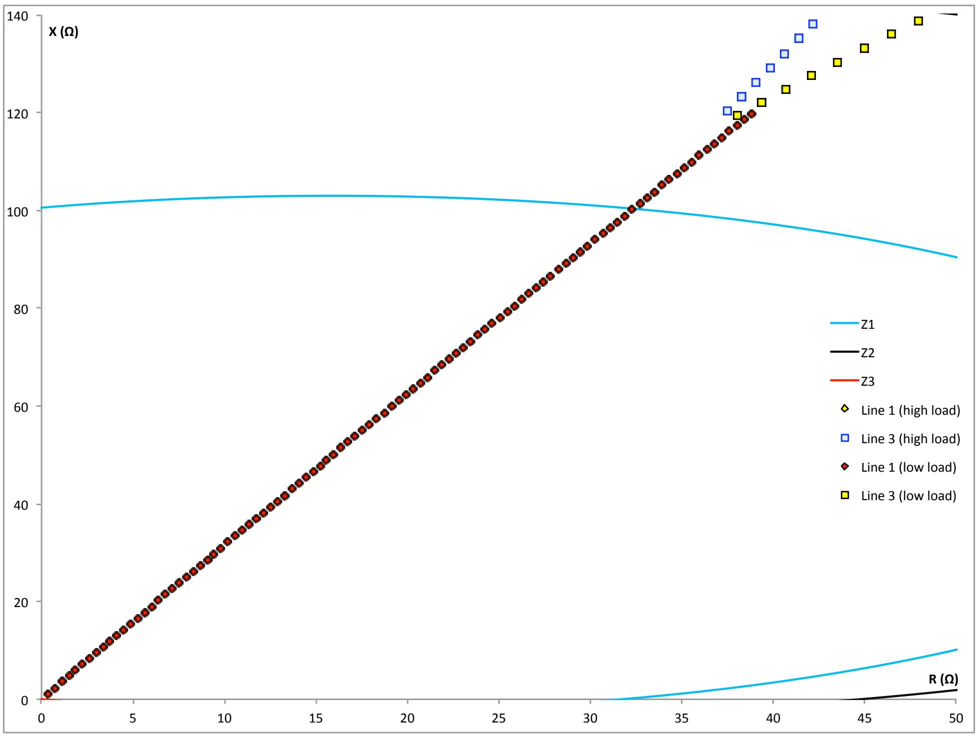

Figure 3.7: In-feed effect (detail).

The load condition influence on the relay impedance measurement presents different behaviours that depend on the line and distance of the fault sampled. To analyse this behaviour, 3-phase faults at 5% and 95% of each line are simulated considering different pre-fault conditions (each fault location simulated includes 100 samples). The fault locations are presented in figure 3.8.

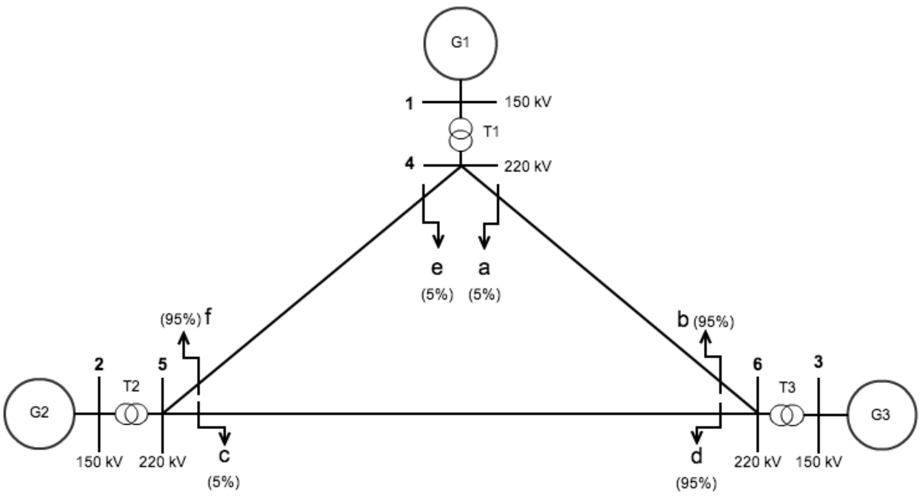

Figure 3.8: Fault Locations.

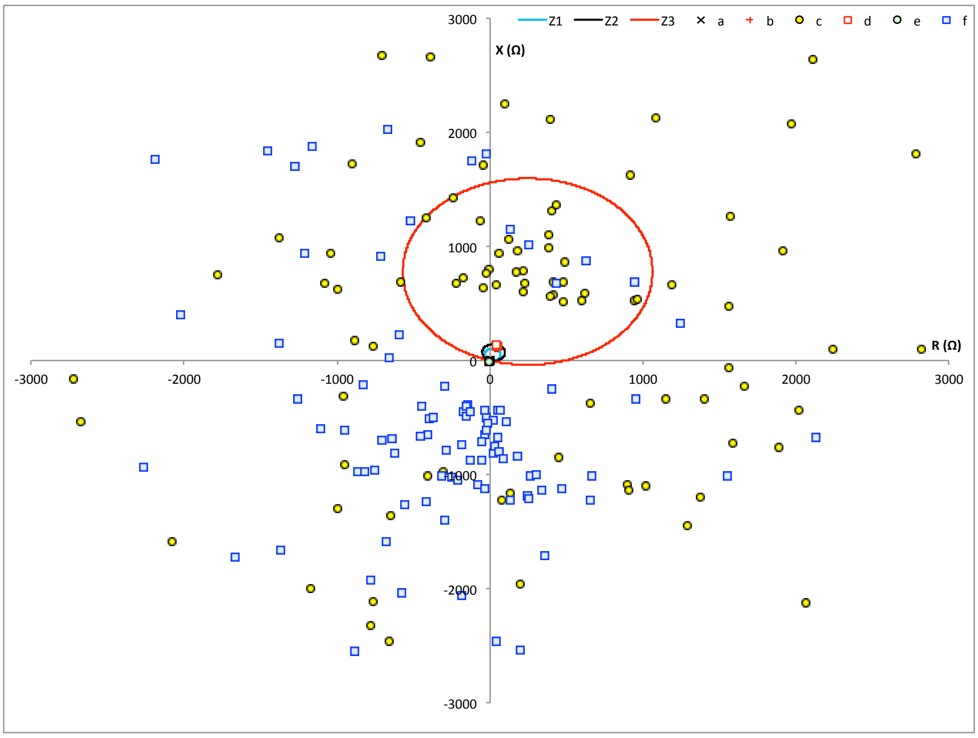

Figure 3.9: Pre-fault load condition effect.

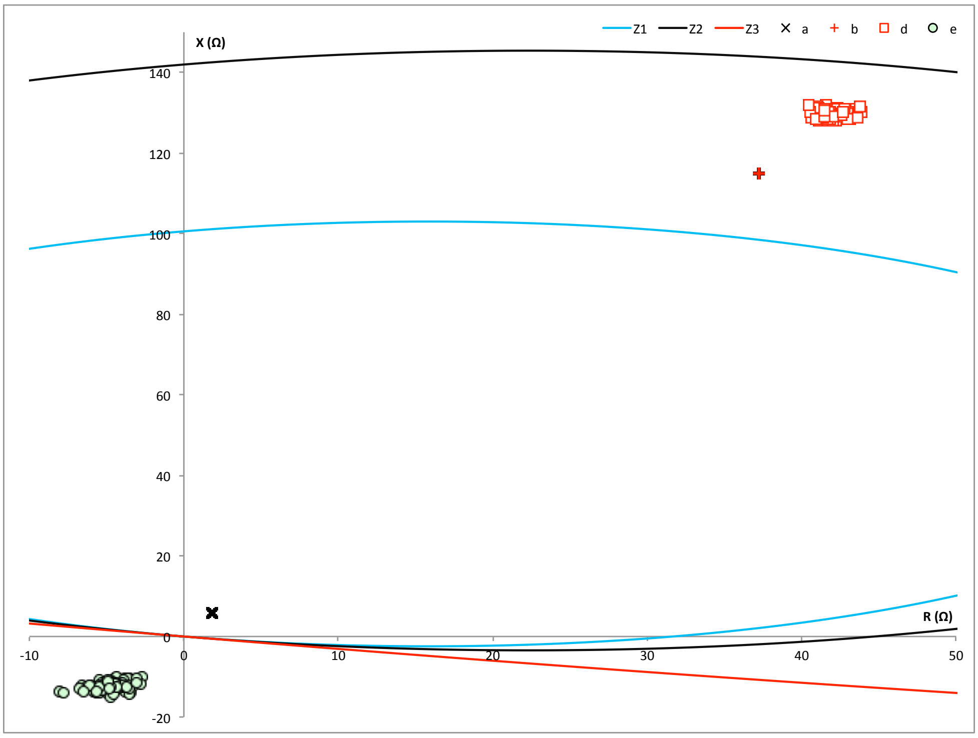

Figure 3.10: Pre-fault load condition (detail).

3.2.3 Analysis of the Impedance Behaviour

From figures 3.10 and 3.7, and equation 3.1 it can be seen that the impedance measurement for faults in line 1 is not influenced by any of the effects tested. On the other hand, from figures 3.9 and 3.6, and equations 3.2 and 3.3, the influence of pre-fault condition and intermediate in-feed is clearly noticed in lines 2 and 3. The analysis of this results leads to the following conclusions:

- The intermediate in-feed in a meshed system causes a non-linear deformation in the impedance measured due to the distribution of current flows. This distribution depends on the impedance of the parallel paths separating the generators from the fault: as the fault location varies the impedance of the parallel paths varies causing the current flow in the lines to vary. The variation of the fault location in lines 2 and 3 causes the current flow in line 1 to vary due to the distribution of current in-feeds from generators G1, G2 and G3, which in turn causes the non-linear variation of the fault impedance. This behaviour is verified by the equations deduced in subsection 3.2.1.

- Figure 3.6 shows that as the fault location moves away from the location of the relay the in-feed currents distribution in the lines induces a higher displacement of the measured impedance. This behaviour is also verified by equations 3.2 and 3.3: in lines 2 and 3 the numerator of the function that relates the measured impedance with the fault distance presents a quadratic variation with fault distance, while the denominator presents a linear variation with fault distance. This relation causes the impedance growth rate of the numerator with fault distance to be higher than the numerator impedance growth rate, which means that as the distance increases the variation of the measured impedance also increases.

- The pre-fault load condition only influences the measured impedance if there is an intermediate in-feed between the relay and the fault location. This condition influences the active and reactive current flows due to the existence of different load admittances in buses 4, 5 and 6, which in turn influences the resistance and reactance of the impedance measured.

- Since the pre-fault condition influences the in-feed currents, it can be concluded that for fault locations where these present a stronger effect, the pre-fault condition has a stronger influence in the impedance measurement. The influence of this condition combined with the influence of the in-feeds for faults located in positions c and f is so high that it is capable of altering the direction of the current flow in line 1. Figure 3.9 shows that faults in positions c and f may adopt different active and reactive impedance measurements in both directions.

3.3 Mho Distance Relay Operation Analysis

The final section of this chapter presents an analysis of the mho relay performance. For this purpose a data sample of 30000 power flow conditions and 30000 faults was created, being 9939 faults in line 1, 9847 faults in line 2 and 10214 faults in line 3 (the data points generated are presented in the R-X diagram in D).

The perspective adopted in this evaluation aims at identifying and quantifying the incorrect and correct zone detection situations. To do so, the correct zone identification, determined from the previous knowledge of the system operating condition or fault location, is compared with the zone where the impedance point measured by the relay lies in the R-X diagram. The correct zone identification is based on the following assumptions:

1. **Faults in the the first 85% of line 1**: Zone 1.
2. **Faults in the the last 15% of line 1**: Zone 2.
3. **Faults in the last 38% of line 3**: Zone 2 or 3.
4. **Faults in the first 62% of line 3**: Zone 3.
5. **Faults in line 2 or normal operation**: Zone 3.

All fault locations present a specific zone, except item 3. The reason behind this assumption lies in the pre-fault load flow effect: zone 2 reach is dimensioned as 120% of the impedance of line 1, so the percentage of line 3 covered by zone 2 can't be precisely determined. If a certain power flow is assumed, this number could be correctly calculated for the case considered, but the option for a certain pre-fault condition introduces an error, because a correct pre-fault condition does not exist. To solve this question a maximum zone 2 reach is assumed. In subsection 2.3.2 is written that the maximum zone 2 reach is equal to:

$$Z2_{maximumreach} = 0.80 * [Z_{line1} + 0.85 * Z_{line3}] \quad (3.4)$$

$$Z2_{maximumreach} = 46.6992 + j148.2944 \ (\Omega) \quad (3.5)$$

This value is equivalent to the sum of the impedance of line 1 plus the impedance of the last 38% of line 3, so it is assumed that the maximum reach of zone 2 in line 3 is 38%. With this approach the problem of assuming a specific pre-fault load flow is overcome while respecting the rules of relay setting.

To evaluate the zone classification performance of the mho relay a 4x4 table is adopted considering all the possible correct and incorrect situations. The diagonal represents the cases where the zone detected is correct while the other cells contain the incorrect classifications:

Table 3.2: Mho relay performance on zone detection.

		Zone detected by the mho relay			
		Out of protected zone	Zone 1	Zone 2	Zone 3
Correct zone	Out of protected zone	39756	0	0	91
	Zone 1	0	8505	0	0
	Zone 2	0	0	2410	0
	Zone 3	1062	0	0	8176

The results presented in table 3.2 show that the relay can discriminate all faults occurring in zone 1 and zone 2, but can not discriminate correctly two situations:

1. **Correct zone = Out of protected zone - Zone detected by the relay = Zone 3**.
   The cases quantified in this cell include faults in line 2 detected in zone 3 (39 cases with distances between 93.2562022% and 99.9400201%) and heavy load conditions (53 cases).
2. **Correct zone = Zone 3 - Zone detected by the relay = Out of protected zone**.
   The cases quantified in this cell consist in faults in line 3 that the relay detects out of the protected zone 3 (1062 cases with distances between 0.00702607% and 28.5547753%) .

This behaviour is expected since the impedance measurement for faults in lines 2 and 3 is affected by pre-fault condition and intermediate in-feed. The difficulties are in discriminating heavy load conditions from faults and identifying faults occurring near the boundary between lines 2 and 3, being line 3 the one that presents the higher error in number of cases and distance. This error could be reduced by increasing zone 3 reach which, one the other hand, would increase the number of faults in line 2 detected in line 3 and heavy load impedance points detected in zone 3.

The results obtained show that the relay operates incorrectly in zone 3.

Chapter 4

Improving the Performance in Zone 3

4.1 On the Importance of Zone 3

Zone 3 is a time delayed overreaching distance zone with the objective of offering remote back-up protection to Zone 1 and Zone 2 of adjacent lines in the case of relay or breaker failure. Recent relay packages include alternatives to offer back-up protection to relay or breaker failure; namely the inclusion of two pilot relays or breaker failure protection schemes; reducing the possibility of failure and the need for a zone 3 element. The necessity of Zone 3 varies with system topology and protection requirements, for each of the following contingencies the importance of a Zone 3 element is evaluated [7]:

- **Batteries**: a failure in the battery responsible for feeding the relay.
  When the relay is installed in substations with a SCADA system an alert signal is issued in case of battery malfunction so that the necessary replacement actions are executed. While the relay is not operative a remote back-up protection is necessary. At higher voltage levels is normal to have a second battery which reduces the time spent in the replacement, thereby reducing the need for a zone 3 element.

- **Relays**: intern relay failure.
  In the case of higher voltage levels two pilot relays are installed simultaneously. This measure reduces the probability of failure, therefore the necessity for a back-up zone is reduced. Otherwise, the existence of a zone 3 element is indispensable.

- **Transducers**: incipient faults in current and voltage transformers.
  At higher voltage levels these equipments are duplicated or fused separately which reduces the probability of error. In case of lower voltage levels these equipments are not duplicated, making the existence of back-up protection a necessity.

- **Circuit Breakers**: the circuit breaker does not open when the trip order is issued.
  These equipments can not be duplicated, so to clear a fault a transfer trip scheme or a

zone 3 element must be adopted. The first option involves costs with communication equipment, which may be unnecessary, making the adoption of a zone 3 element a better choice.

- **Catastrophic Station Failure**: failure of the relay and back-up equipment due to natural causes or human error (incorrect setting or equipment outage during maintenance). In these cases the only option is the adoption of a zone 3 element.

From this analysis it may be concluded that the existence of communication channels and duplicate devices reduces the necessity for a zone 3 element, nonetheless this zone always represents a remote back-up protection that increases the reliability of the system and is an indispensable function when these options are not available.

4.2 Towards a Neural Network Fault Locator

The results of the study conducted in chapter 3 showed that for the system topology selected the relay operates incorrectly in zone 3 in two different cases: for faults in the boundary of line 2 and 3, and for heavy load conditions. To improve the number of correctly identified faults in line 3 the reach of zone 3 must be increased, resulting in an increase in the number of unintended trip situations. On the other hand, to reduce the number of unintended trip situations the reach of zone 3 must be reduced and the number of faults in line 3 that will not be detected increases. In this case the protection engineer is facing a multi-criteria decision problem that implies determining the ideal compromise between the number of underreaching and overreaching situations. In this dissertation this problem is not addressed, instead a new solution for zone 3 tripping is proposed, inspired on the fault location process.

Fault location algorithms (see [18] , [19], [20],[21]) are used to determine the location of the fault, so that the service restoration time and reliability of the system are improved. This feature consists of two steps: in the first step the type of fault and phase or phases involved are identified, in the last step the distance between the relay location and the fault location in the line being protected is estimated. In this dissertation a scheme inspired in fault location algorithms is applied to improve the operation of zone 3 (since this zone is normally dimensioned for a tripping time of 850 ms [4], the time elapsed in the calculation process is not a problem). The basic concept includes detecting a fault, identifying its location and determining the zone. This approach aims to optimize the capacity of the relay to correctly identify faults inside its protection zone by the relay without increasing the number of cases of unintended tripping.

To evaluate the performance of this concept the data set presented in appendix D is used. In this data set only three phase faults were sampled so the fault classification process is not applied as a fault type identifier, instead a line classification approach is adopted: since the fault presents different behaviours for the different lines it is necessary to identify the line to obtain a correct distance estimation. In short, the solution being proposed consists in the following steps:

1. **Fault detection**: discriminate normal operation from fault conditions.
2. **Line Identification**: determine the faulted line.
3. **Distance estimation**: estimate the location of the fault on the line affected.
4. **Zone identification**: determine the zone where the fault occurs.

Analysing the data in appendix D it can be observed that the problem of line identification involves the definition of non-linear and disjoint zones, which make the mho characteristic an inadequate zone shape. This behaviour, has stated in chapter 3, is related to the effect of intermediate in-feeds and pre-fault condition, so to obtain a better classification performance the solution proposed must be able to define non-linear boundaries and adapt itself to the pre-fault operating condition. In respect to the distance estimation process the solution must be able of determining the location of the fault in the line from the relationship between the measured voltage and current signals at the relay location. In section 3.2 it is shown that the determination of this value for faults in line 1 is a simple procedure and possible to be executed with a simple mathematical demonstration, for lines 2 and 3 to obtain the distance of the fault in the line the in-feeds from the different sources must be known. Since communication links are not considered, the solution must incorporate a method capable of determining this value from the measured signals at the relay location. Considering the guidelines referred the inputs and method applied are:

- **Input**
  The digital conversion of the measured voltage and current signals is performed by the relay at a sampling rate of 16 samples per fundamental power cycle, corresponding to a frequency of 960Hz [15], in the solution proposed the input consists of a data window of two sequential measures, namely the voltage and current signals and the angle between them, measured at the operating moment and these same values measured in the sample 1.25 ms before. With this approach when a fault occurs the value measured before is related to the pre-fault operating condition and the value measured in the moment is related to the sub-transitory short-circuit condition. To simulate normal operating conditions it is considered that in the time interval of 1.25 ms, if no disturbance occurs, the measures in these two moments are equal. During operation these values are continuously injected in a pipeline mode: when new measured values enter the ones that occurred 2.5 ms before leave.

- **Neural Networks in fault detection, line classification and distance estimation**
  Neural networks are a computational intelligence technique. These are multilayered parallel structures that acquire knowledge through a process of "learning", storing it in

the form of weights and forces of connection between neurons. The reasons behind the option for the application of neural networks lie in the following characteristics[8]:

– **Non-Linearity**: the aptitude of recognizing non-linear patterns and reproducing non-linear functions that cannot be modelled by mathematical expressions.

– **Mapping Input-Output**: the capacity of classifying patterns in a more precise way.

– **Generalization**: the ability to generate an output to entries that have not been presented during the training (the training process is executed off-line and it is based on simulation results, since it is not possible to incorporate all the possible situations during the training these ability is vital).

The fault detection and line classification processes are incorporated in the same classification neural network. The reason behind the junction of these two steps is that both of them consist of a classification process, besides it was observed through experimentation that this junction does not influence the performance of neither of the steps. The distance estimation consists of a function approximation neural network for each of the lines in the system.

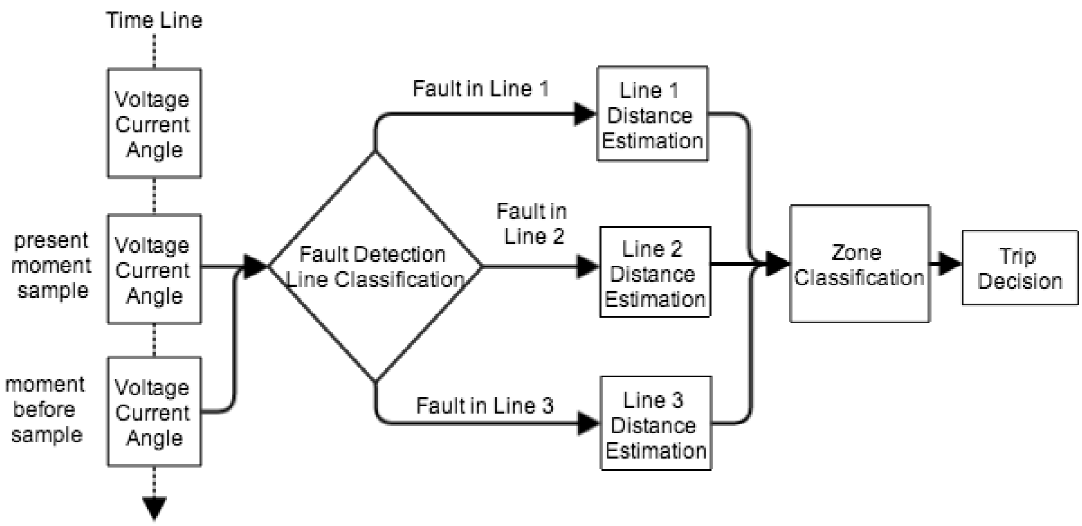

Figure 4.1: Architecture of the solution.

This solution (see figure 4.1), besides its application in the improvement of zone 3 performance, can be used for a possible reverse zone 3 and to optimize the post-fault location analysis, reducing the time spent in fault location and thereby improving the reliability of the system and reducing the costs associated.

4.3 Modelling the Neural Network Fault Locator

The steps presented in the solution proposed include 2 different problems to be solved by neural networks: pattern recognition (in fault detection and line identification) and function approximation (in distance estimation).

Neural networks for pattern recognition intend to reproduce non-mechanical tasks, such as perception, memory and conscious thought, in description, classification and grouping of patterns. These techniques have been used in different areas from document classification, diagnosis, data mining, financial forecasting among others [33]. In this dissertation a classification problem is addressed using neural networks following two distinct approaches: the first considers a feedforward neural network (FNN) and the second a structure of competitive autoencoders.

Neural networks for function approximation intend to map the output values of a given set of input values. This is a supervised training process in which an unknown function, that reproduces a certain output to a given input, is simulated by the neural network [34]. In this dissertation a FNN is used for function approximation.

4.3.1 Feedforward Neural Network for Classification

A neural network is said to be feedforward when there are only connections between successive layers, in other words, there is no feed-back (this type of neural network is capable of simulating deterministic functions and executing multivariate non-linear functional mapping[6]).

The classification problem in FNNs may be solved by associating a specific codification to each class. During training, different instances belonging to different classes are presented as input, and the desired output is the codification associated with the class of the instance presented. The objective function is the mean square error between the output obtained and the respective target, which is minimized by the training algorithm. The output obtained is not an integer so this value is rounded to obtain a classification. This process requires the definition of non-linear multivariate boundaries to separate the different classes, making the option for FNNs an adequate choice.

To define the adequate structure, the capacity of FNNs to generate non-linear boundaries was studied in terms of number of layers, type of activation functions and layer dimension:

- **Number of Layers**

  FNNs with a different number layers can produce different decision boundaries [6]. A possible representation of this relation is presented in figure 4.2.

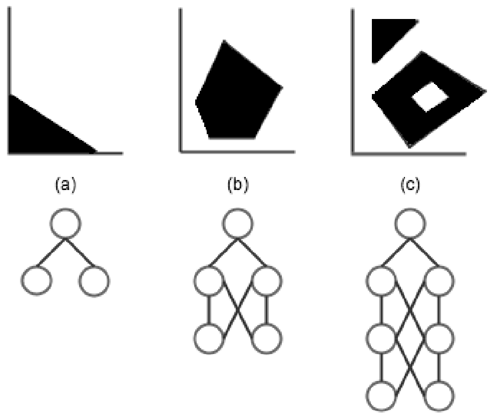

Figure 4.2: Possible decision boundaries which can be generated by different numbers of layers in a feedforward neural network ([6]).

  In figure 4.2 it can be seen that different number of layers restrict the complexity of the boundaries which delimit the different classes. For a neural network with one layer of weights (a) the decision boundary is a hyperplane. For the case of two layers of weights (b) a single convex region of the input space delimited by segments of hyperplanes may be created. Finally, for three layers of weights (c) the neural network is capable of defining arbitrary decision regions, non-convex and disjoint [6].

- **Activation Functions**
  Neural networks composed by linear activation functions are only capable of defining hyperplanes. Since each layer represents a transformation in a multivariate space a layer with linear activation functions performs a linear transformation. The introduction of non-linear activation functions produces non-linear transformations which are able to define non-linear boundaries[6].

- **Layer Dimension**
  Neural networks with a lower number of neurons in the hidden layers produce a data compression which leads to the loss of information[6], an adequate approach for problems with redundant information. On the other hand, if the number of neurons in the hidden layer is increased, noise may be introduced which complicates the process of classification.

Due to the factors stated and the analysis of the data set, the FFN used in the fault detection and line classification block is composed by four layers of neurons, where the first three layers have the same number of neurons and the fourth layer is composed of a single neuron (see figure 4.3). The inclusion of non-linear activation functions must take place to introduce the possibility of defining non-linear boundaries. The neural network is trained to present as output the code number that represents the class of a certain input instance, being these classes codified by:

- Normal operation: -1.
- Fault in Line 1: 1.
- Fault in Line 2: 2.
- Fault in Line 3: 3.

The state associated to a certain instance is determined by rounding the output value of the neural network (e.g: if the output of a certain instance is one than the correspondent state is fault in line 1).

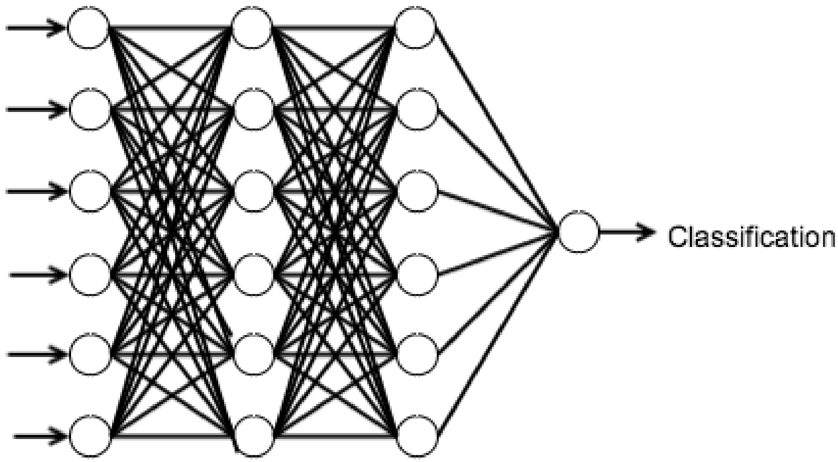

Figure 4.3: Feedforward neural network topology used for fault detection and line classification.

4.3.2 Competitive Autoencoders for Classification

Autoencoders or autoassociative neural networks are FNNs trained to reproduce the input vector in the output, which means that the number of inputs and outputs is the same. The most simple autoencoder topology is composed by only one hidden layer, although there is no theoretical limit to the dimensionality of the autoencoder.

A typically used topology consists in the adoption of three layers, namely the input layer, the middle layer and output layer, being the middle layer composed by a lower number of neurons (see figure 4.4). In this architecture the input vector is projected into a compressed space, present in the middle layer, and afterwards is expanded to the original space, in the output layer, through the application of the inverse transformation. This type of neural network is commonly used in data compression ([35], [36]) and missing data restoration ([37], [38]).

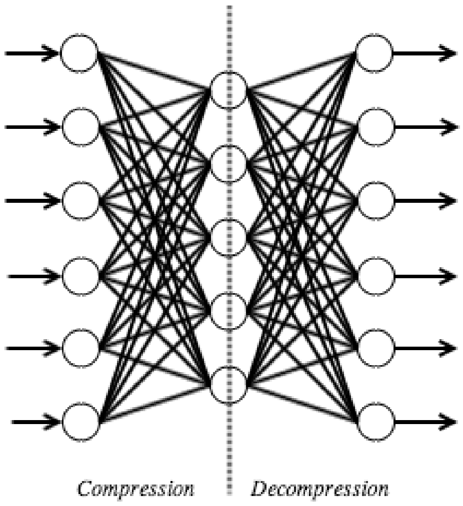

Figure 4.4: Bottleneck Autoencoder.

An autoencoder is trained as any other kind of FNN. The most common objective function is the mean square error, which is minimized through the backpropagation of the error. For the training process three data sets are used: one for training, one for testing and one for validation.

An interesting capacity of autoencoders with linear activation functions is that they reproduce a mapping on the hidden layer similar to a Principal Component Analysis [39]. With this technique the information is condensed in the orthogonal axes so that variance is minimized. If non-linear activation functions are applied the mapping is not equivalent to PCA, since the axes present a non-linear shape, and a better mapping is obtained for problems with a non-linear behaviours [40].

Once trained, an autoencoder is capable of discriminating instances belonging to a certain training set from instances presenting different characteristics. If the instance presented at the input belongs to the manifold obtained with the training data set, the error between the input and output values will be low. On the other hand, if the input instance contains different characteristics, it will not be correctly re-projected and the error between the input and output vectors will be large. This characteristic has been explored in [41] and [42] through the creation of a competitive structure of autoencoders for pattern recognition.

The idea behind the creation of a competitive autoencoder structure is simply training different autoenconders to recognize different data sets and then classify a certain unknown instance by determining the autoencoder which presents the minor error. The set represented by this autoencoder is considered to be the one to which the unknown instance belongs. With this approach the instance is classified by its resemblance to the training set of each autoencoder.

The application of this method in the fault detection and line classification block is made by training four autoencoders: one is trained to recognize the normal operation state and the other three represent faults in each of the three lines (see figure 4.5). The state associated with a certain sample is determined by the auto-encoder which presents the minor mean square error to that specific sample.

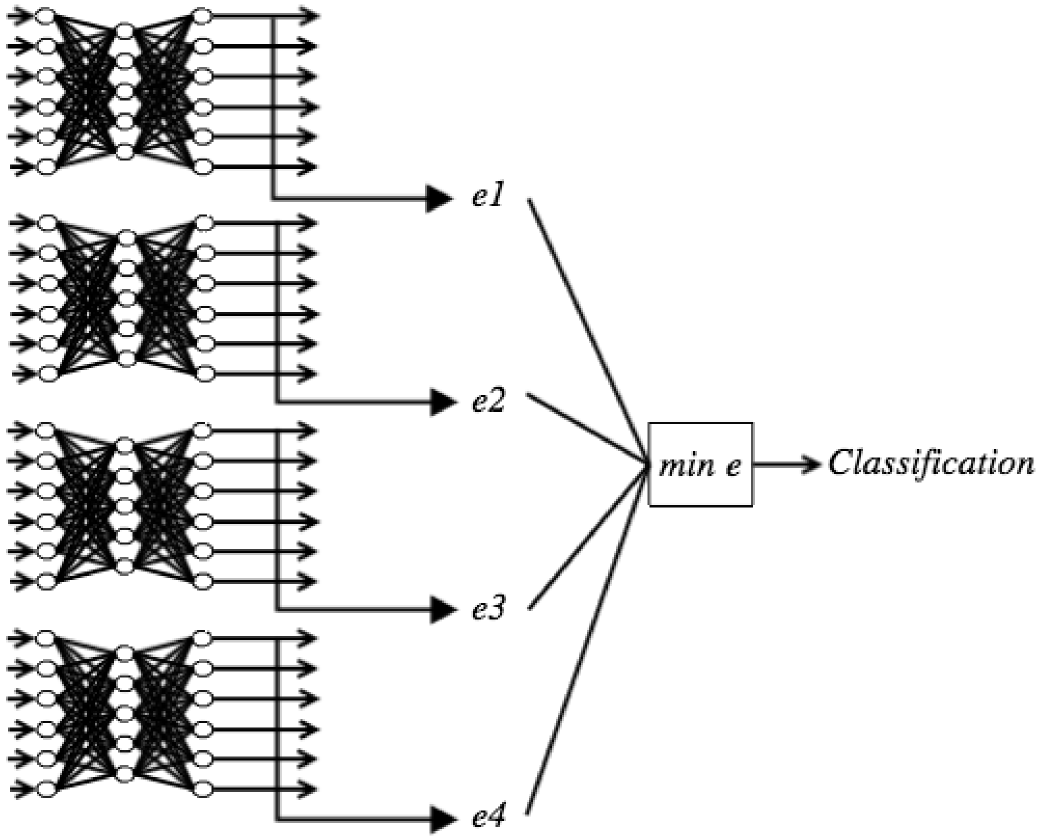

Figure 4.5: Competitive autoencoders structure used for fault detection and line classification.

4.3.3 Feedforward Neural Network for Function Approximation

FNNs, when applied to function approximation, have the objective of reproducing the unknown function behind the relation between the pairs of inputs and outputs of a certain manifold. To learn this relation the minimization of the error between the machine output and the desired response is applied as the objective function. Through this procedure the machine approximates the unknown function with the input - output mapping of the data manifold [34].

The simplest way of approximating a function through a FNN is by using the projection theorem principles. This theorem states that the behaviour of a certain function $f(x)$ can be mimicked in a certain input space by a combination of simpler functions φ_i [34]:

$$F(x,w) = \Sigma^{N}{}_{i} * w_i * \varphi_i(x) \tag{4.1}$$

where x is the input, w_i the entries of the coefficient vector and N the number of neurons in the hidden layer. The neural network is trained such that:

$$min\ \varepsilon = |f(x) - F(x,w)| \tag{4.2}$$

The number of simpler functions used depends on the error ε desired. As the number of simpler functions combined increases the ability of the neural network to adapt to a certain shape also increases[34].

The FNN used for the fault distance estimation of each line is based on this theorem and consists of two layers: the hidden layer and the output layer. For each set of inputs the desired output is the fault distance. The topology of the neural network used is represented in figure 4.6. The number of neurons in the middle layer is defined by a trial and error process.

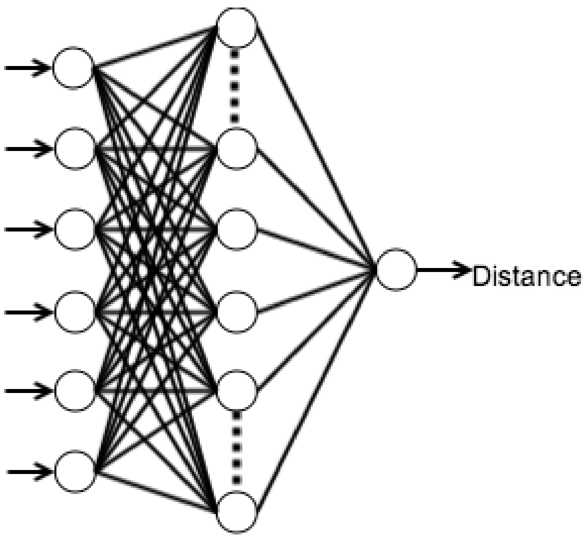

Figure 4.6: Feedforward neural network used for fault distance estimation.

4.4 Performance Evaluation

In this section the results obtained in the tests performed with the methods proposed are presented. The Levenberg-Marquardt training algorithm is used in every test. The neural networks were created using the MATLAB neural network toolbox [43]. The data set used for training, test and validation is presented is composed by 9939 cases of faults in line 1, 9847 cases of faults in line 2, 10214 cases of faults in line 3 and 30000 cases of normal operation (see appendix D). Depending on the neural network to be trained the respective inputs are selected and divided, being 70% of the set for training, 15% for validation and 15% for test. The performance of the neural network fault locator is evaluated by analysing the correct responses for all the cases sampled.

The neural network training procedure is based on an iterative method, so the evolution to the optimum depends on the initial point. In the case of neural networks this point includes the weight and bias initialization. For weight initialization the transformation matrix of the principal component analysis (PCA) applied to the data set used for training, test and validation is used, since through experimentation it proved to be a good initialization. For bias initialization the same value is applied to each layer, being it 0 or 1. The values adopted are the ones that through a trial and error procedure proved to achieve the best result.

4.4.1 Fault Detection and Line Classification

To evaluate the performance of the two approaches proposed for this function block a 4x4 table containing the correct state and the state determined by the neural network is presented for each case. For comparison purposes the results of a scheme only considering as inputs the voltage, current and angle values measured in the operating moment is presented.

4.4.1.1 Feedforward Neural Network

- First Middle Layer Activation Function: hyperbolic tangent.
- Second Middle Layer Activation Function: hyperbolic tangent.
- Output Layer Activation Function: linear.

Table 4.1: Fault detection and line classification using a feedforward neural network.

		State detected by the neural network			
		Normal Operation	Fault Line 1	Fault Line 2	Fault Line 3
Correct state	Normal Operation	30000	0	0	0
	Fault Line 1	0	9939	0	0
	Fault Line 2	0	0	9756	91
	Fault Line 3	0	0	129	10085

The results presented in table 4.1 show that the neural network can discriminate normal operation conditions and faults in line 1, but can not discriminate correctly two situations:

1. **Correct state = Line 2 - State detected by the neural network = Line 3**.
   The cases quantified in this cell are composed by faults in line 2 detected in line 3 (91 cases with distances between 74.8874651% and 99.8976311%).
2. **Correct state = Line 3 - State detected by the neural network = Line 2**.
   The cases quantified in this cell consist in faults in line 3 that the neural network detects as being in line 2 (129 cases with distances between 0.0417638% and 14.0048644%) .

Considering three inputs, namely the voltage, current and angle measured at the present moment, the neural network used is composed by only three inputs and six neurons in the middle layers. The activation functions are equal to the ones used in the six inputs case.

Table 4.2: Fault detection and line classification using a feedforward neural network considering only three inputs.

		State detected by the neural network			
		Normal Operation	Fault Line 1	Fault Line 2	Fault Line 3
Correct state	Normal Operation	30000	0	0	0
	Fault Line 1	0	9939	0	0
	Fault Line 2	0	0	9442	393
	Fault Line 3	0	0	397	9867

The results presented in table 4.2 show that the neural network with three inputs presents a similar behaviour when compared to the approach with six inputs but with a larger error:

1. **Correct state = Line 2 - State detected by the neural network = Line 3**.
   The cases quantified in this cell are composed by faults in line 2 detected in line 3 (393 cases with distances between 77.6993178% and 99.9373521%).
2. **Correct state = Line 3 - State detected by the neural network = Line 2**.
   The cases quantified in this cell consist of faults in line 3 that the neural network detects as being in line 2 (397 cases with distances between 0.0417638% and 30.7749216%) .

4.4.1.2 Competitive Autoencoders

- Middle Layer Activation Function: hyperbolic tangent.
- Output Layer Activation Function: linear.

Table 4.3: Fault detection and line classification using competitive autoencoders.

		State detected by the competitive autoencoders			
		Normal Operation	Fault Line 1	Fault Line 2	Fault Line 3
Correct state	Normal Operation	30000	0	0	0
	Fault Line 1	0	9939	0	0
	Fault Line 2	0	1507	8330	10
	Fault Line 3	0	3839	1260	5115

The results presented in table 4.3 show that with this method is possible to discriminate fault conditions from normal operation, but the faulted line can not be correctly determined. To study this behaviour the data samples of faults in the different lines was injected in each autoencoder and the mean error was calculated for every case:

Table 4.4: Mean absolute error of the autoencoder for the different line faults data samples.

		Autoencoders		
		Line 1	Line 2	Line 3
Sample	Line 1	2.41E-08	0.080239853	0.036625276
	Line 2	0.064527093	0.000150752	0.035571205
	Line 3	0.054284542	0.009526055	0.000389699

4.4.2 Fault Distance Estimation

Test were conducted to determine the mean absolute error of distance estimation for each line considering different numbers of neurons in the hidden layer. The activation functions of the distance estimation neural networks are:

- Middle Layer Activation Function: hyperbolic tangent.
- Output Layer Activation Function: linear.

The evolution of the estimation error with the growing number of neurons in the hidden layer for each neural network are presented in figures 4.7 and 4.8. The stopping criteria used in the process of training is 1000 epochs.

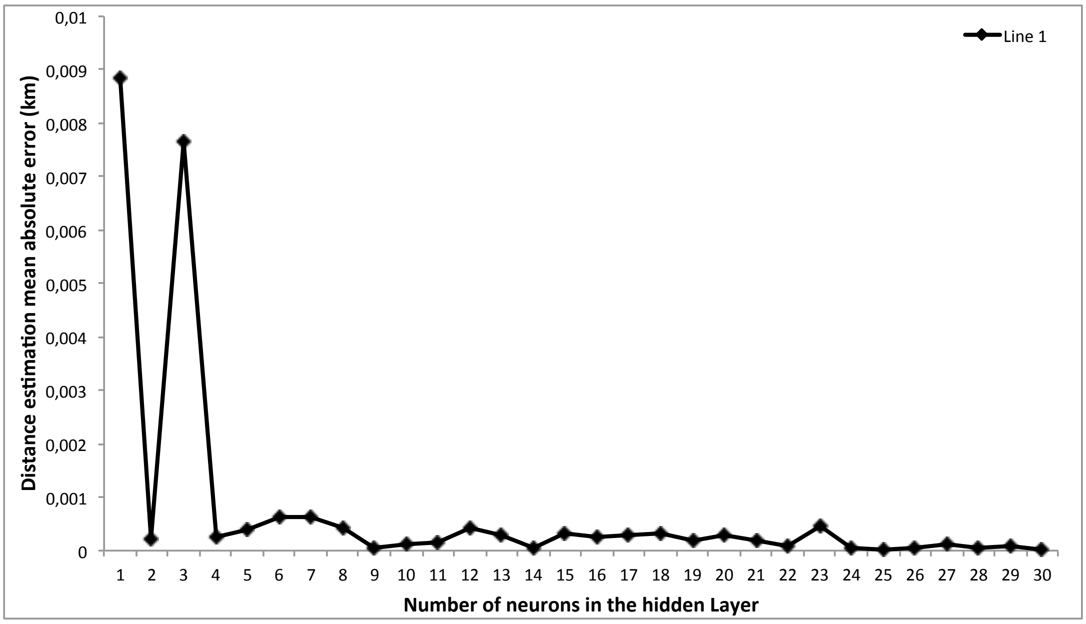

Figure 4.7: Line 1 neural network distance estimation mean absolute error for different number of neurons in the hidden layer.

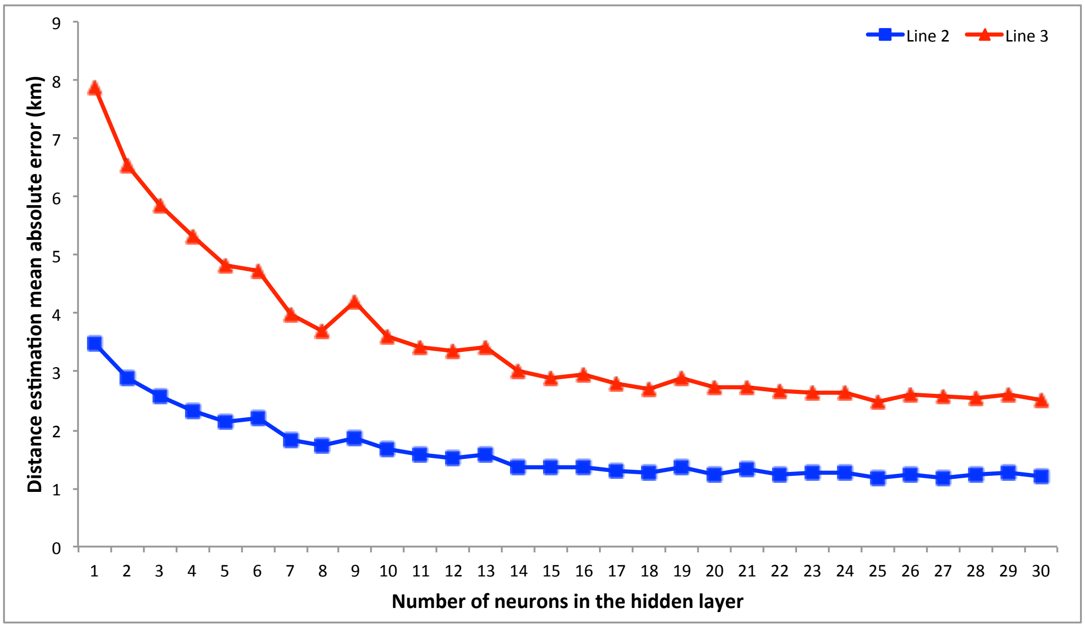

Figure 4.8: Lines 2 and 3 neural networks distance estimation mean absolute errors for different number of neurons in the hidden layer.

4.4.3 Analysis of the Tests Performed

- Table 4.1 shows a better line identification performance than table 4.2. This result shows that the neural network is able to learn the relation between the pre-fault load condition and the fault condition in the relay inputs.

- Comparing tables 4.1 and 4.2 with table 4.3 it can be observed that one FNN with three layers of weights achieves a better performance in pattern recognition than the competitive autoencoders structure in the problem in study. Table 4.4 shows that the autoencoder trained to recognize line 1 faults presents a very low error for faults in line 1 when compared to faults in the other lines. This result is not observed for the autoencoders trained to recognize faults in line 2 and 3, although the error corresponding to the data set for which these autoencoders where trained is lower than the error corresponding to the data samples of faults in the other lines, this difference is not high enough to achieve a correct line classification.

- In figures 4.7 and 4.8 it can be observed that the fault distance is determined with a very low error for faults in line 1 when compared to faults in lines 2 and 3. This result is expected since in line 1 the relation between the distance and the impedance measured by the relay is linear and can be exactly determined by dividing the voltage and current signals measured by the relay (see equation 3.1), while in lines 2 and 3 this relation is non-linear and can not be determined by the relay with only the signals measured at the relay location (see equations 3.2 and (3.3)).

- In figure 4.8 it can be observed that as the number of neurons in the hidden layer grows the error of distance estimation reduces, this behaviour shows that for a high number of neurons in the hidden layer the accuracy of the distance estimation may be greatly increased.

4.5 Final Solution

The final solution consists in the combination of the neural network topologies which present the best performance. The final solution is composed by the following neural networks:

- Fault Detection and Line Classification: FFN with six inputs.
- Line 1 Distance Estimation: FNN with 2 neurons in the hidden layer.
- Line 2 Distance Estimation: FNN with 18 neurons in the hidden layer.
- Line 3 Distance Estimation: FNN with 18 neurons in the hidden layer.

The number of neurons in the hidden layers of the distance estimation blocks is selected so that a compromise is achieved between the dimension of the hidden layer and the distance estimation accuracy. Line 1 presents a high accuracy with only two neurons so this topology is adopted. In the case of lines 2 and 3 it can be seen that with two neurons the error is very high so a number of 18 neurons is adopted to obtain a acceptable accuracy. The result obtained in zone classification with these combination of neural networks is:

Table 4.5: Neural network fault locator performance on zone classification.

		Zone detected by the neural network fault locator			
		Out of protected zone	Zone 1	Zone 2	Zone 3
Correct zone	Out of protected zone	39756	0	0	91
	Zone 1	0	8505	0	0
	Zone 2	0	0	2478	0
	Zone 3	129	0	0	9041

Analysing tables 4.1 and 4.5 it can be observed that the incorrect zone classification cases are the same as the incorrect line classification cases. This result shows that the error in zone classification for the validation set used is only dependent in the performance of the fault detection and line classification block, so to optimize the performance of the solution proposed a better performance must be achieved in pattern recognition.

Remembering the the performance of the mho relay:

Table 4.6: Mho relay performance on zone classification.

		Zone detected by the mho relay			
		Out of protected zone	Zone 1	Zone 2	Zone 3
Correct zone	Out of protected zone	39756	0	0	91
	Zone 1	0	8505	0	0
	Zone 2	0	0	2410	0
	Zone 3	1062	0	0	8176

Comparing tables 4.6 and 4.5 it is observed that the neural network fault locator obtains a higher number of correct decisions and improves the performance of zone 3. The solution proposed optimizes the following cases: the incorrect trip decisions due to heavy load are eliminated and a higher number of line 3 faults are correctly identified (933 cases). The number of faults in line 2 identified as zone 3 to increase in 53 cases.

Chapter 5

Conclusion

5.1 General Conclusions

In this dissertation the effects of pre-fault load condition and in-feeds on the behaviour of the impedance measured by the distance relay were studied, besides the errors on the distance protection performance were identified and a solution based on neural networks was proposed. The main conclusions of this work are divided in two themes:

Measured Impedance Behaviour And Relay Operating Errors

1. Intermediate in-feeds on meshed system configurations cause non-linear variations on the impedance measured by the relay.
2. Pre-fault load condition only affects the impedance measured by the relay for faults in lines with intermediate in-feeds.
3. The admittances of the loads from the pre-fault load condition influence the current flow during a short-circuit event and may cause the impedance measured by the relay to modify greatly.
4. The combination of intermediate in-feed and the variety of pre-fault conditions cause the mho relay to operate incorrectly in zone 3 for the validation set tested.
5. Heavy load conditions cause the relay to operate incorrectly in zone 3 for the validation set tested.

Neural Networks Improvement In Distance Protection Performance.

1. Neural networks performing pattern recognition are efficient in discriminating non fault from fault conditions (100% of correct decisions for the validation set tested) with only the voltage and current signals measured at the relay location.
2. Neural networks are capable of learning the relation between the measured values of voltage, current and angle in the pre-fault loading and short-circuit conditions.

3. The scheme developed presents an affective option for improving the accuracy of zone 3.
4. Neural Networks may be an option to improve the performance of the distance protection.

5.2 Future Studies and Developments

The solution proposed in this document has a very limited scope, therefore the next step is to introduce complexity in the problem. This can be made by introducing different types of faults, fault resistances, arc resistances, fault inception angles and variable source impedances in the data set sampled. The introduction of this elements requires the insertion of a fault classification block in the fault location scheme which increases the complexity of the problem but I believe that the neural networks will be able to accomplish a similar performance.

An important future study involves the validation of the solution proposed. To accomplish this objective the neural network fault locator scheme developed must be tested in other systems and compared with different types of relay operating characteristics. It must also be tested in a support that incorporates the dynamic behaviour of the test system.

Another interesting future study is the application of the scheme developed for a centralized fault location function to diminish service restoration time and increase the reliability of the system. The idea that I propose to be explored is the development of a centralized feature that determines the location of the fault in a transmission system with the measured signals of voltage and current in the nodes of the transmission system for the sub-transitory short-circuit condition. The basis for defending this idea is the efficiency of neural networks in discriminating non fault from fault conditions, determinate the faulted and estimate the distance to the fault in the line.

References

[1] Gerhard Ziegler. *Numerical Distance Protection - Principles and Applications*. Third edit edition, 2008.

[2] P. M. Anderson. *Power System Protection*. 1999.

[3] Alstom. *MICOMho P446 Fast Multifunction Distance Protection Relay*. Alstom, 2011.

[4] AREVA. *Network Protection & Automation Guide*. Alstom Grid, 2011.

[5] Power System Relay Commitee Working Group. Application of Overreaching Distance Relays, 2009.

[6] Christopher M. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press - Oxford, 1995.

[7] S. H. Horowitz and A. G. Phadke. *Power System Relaying*. John Wiley and Sons, third edition, 2008.

[8] S. S. Haykin. *Neural Networks - A Comprehensive Foundation*. Pearson Prentice Hall, Delhi, India, second edition, 1999.

[9] ABB. *Line distance protection REL670 ANSI*. ABB, 1.2 edition, 2012.

[10] G. E. Digital Energy. *D90Plus Line Distance Protection*. GE Digital Energy, 1.8 edition, 2012.

[11] K. K. Li and L. L. Lai. Ideal operating region of digital distance relay under high resistance earth fault. *Electric Power Systems Research*, 43:215–219, 1997.

[12] B. Das and J. V. Reddy. Fuzzy-Logic-Based Fault Classification Scheme for Digital Distance Protection. *IEEE Transactions on Power Delivery*, 20(2):609–616, April 2005.

[13] S. Vasilic and M. Kezunovic. Fuzzy ART neural network algorithm for classifying the power system faults. *IEEE Transactions on Power Delivery*, 20(2):1306–1314, 2005.

[14] S. R. Samantaray, P. K. Dash, and G. Panda. Fault classification and ground detection using support vector machine. *TENCON 2006*, 00:0–2, 2006.

[15] A. H. Osman, T. Abdelazim, and O. P. Malik. Transmission Line Distance Relaying Using On-Line Trained Neural Networks. *IEEE Transactions on Power Delivery*, 20(2):1257–1264, 2005.

[16] M. Gilany, A. Al-Kandari, and B. Hassan. ANN based technique for enhancement of distance relay performance against open-conductor in HV transmission lines. *2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE)*, 5:50–54, 2010.

[17] H. Meyar-Naimi. A new fuzzy fault locator for series compensated transmission lines. *11th International Conference on Environment and Electrical Engineering*, pages 53–58, 2012.

[18] P. K. Dash, A. K. Pradhan, and G. Panda. A novel fuzzy neural network based distance relaying scheme. *IEEE Transactions on Power Delivery*, (3):902–907, July 2000.

[19] P. K. Dash, A. K. Pradhan, and G. Panda. Application of minimal radial basis function neural network to distance protection. *IEEE Transactions on Power Delivery*, 16(1):68–74, 2001.

[20] Z. Lu, T. Y. Ji, Q. H. Wu, and J. Fitch. An adaptive distance relaying algorithm with a morphological fault detector embedded. *2009 IEEE Power & Energy Society General Meeting*, pages 1–8, 2009.

[21] J. R. de Carvalho, D. V. Coury, C. A. Duque, and D. C. Jorge. Development of detection and classification stages for a new distance protection approach based on cumulants and neural networks. *2011 IEEE Power and Energy Society General Meeting*, pages 1–7, 2011.

[22] K. K. Li, L. L. Lai, and A. K. David. Stand Alone Intelligent Digital Distance Relay. *IEEE Transactions on Power Systems*, 15(1):137–142, 2000.

[23] A. K. Pradhan and G. Joós. Adaptive Distance Relay Setting for Lines Connecting Wind Farms. *IEEE Transactions on Energy Conversion*, 22(1):206–213, March 2007.

[24] Bhavesh R. Bhalja and Rudra P. Maheshwari. High-Resistance Faults on Two Terminal Parallel Transmission Line: Analysis, Simulation Studies, and an Adaptive Distance Relaying Scheme. *IEEE Transactions on Power Delivery*, 22(2):801–812, 2007.

[25] J. Upendar, C. P. Gupta, and G. K. Singh. Comprehensive adaptive distance relaying scheme for parallel transmission lines. *IEEE Transactions on Power Delivery*, 26(2):1039–1052, 2011.

[26] M. Jonsson and J. E. Daalder. An adaptive scheme to prevent undesirable distance protection operation during voltage instability. *IEEE Transactions on Power Delivery*, 18(4):1174–1180, 2003.

[27] Chul-hwan Kim, J. Y. Heo, and R. K. Aggarwal. An enhanced zone 3 algorithm of a distance relay using transient components and state diagram. *IEEE Transactions on Power Delivery*, 20(1):39–46, 2005.

[28] S. I. Lim, C. C. Liu, and S. J. Lee. Blocking of zone 3 relays to prevent cascaded events. *IEEE Transactions on Power Systems*, 23(2):747–754, 2008.

[29] P. V. Navalkar and S. A. Soman. Secure remote backup protection of transmission lines using synchrophasors. *IEEE Transactions on Power Delivery*, 26(1):87–96, 2011.

[30] A. H. Osman and O. P. Malik. Transmission line distance protection based on wavelet transform. *IEEE Transactions on Power Delivery*, 19(2):515–523, 2004.

[31] S. M. Brahma. Distance Relay With Out-of-Step Blocking Function Using Wavelet Transform. *IEEE Transactions on Power Delivery*, 22(3):1360–1366, July 2007.

[32] M. Sanaye-pasand and P. Jafarian. An Adaptive Decision Logic to Enhance Distance Protection of Transmission Lines. *IEEE Transactions on Power Delivery*, 26(4):2134–2144, 2011.

[33] J. K. Basu, D. Bhattacharyya, and T. H. Kim. Use of artificial neural network in pattern recognition. *International Journal of Software Engineering and Its Applications*, 4(2):23–34, 2010.

[34] J. C. Principe, N. R. Euliano, and W. C. Lefebvre. *Neural and Adaptive Systems: Fundamentals Through Simulations*. 2000.

[35] B. Golomb and T. Sejnowski. Sex recognition from faces using neural networks. *in A. Murray (ed.), in Applications for Neural Networks*, pages 71–92, 1995.

[36] G. E. Hinton and R. R. Salakhutdinov. Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(July):504–507, 2006.

[37] V. Miranda, J. Krstulovic, H. Keko, C. Moreira, and J. Pereira. Reconstructing Missing Data in State Estimation With Autoencoders. *IEEE Transactions on Power Systems*, 27(2):604–611, May 2012.

[38] S. Narayanan and R. J. Marks. Set constraint discovery: missing sensor data restoration using autoassociative regression machines. In *Proceedings of the 2002 International Joint Conference on Neural Networks*, pages 0–5, 2002.

[39] T. D. Sanger. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, 2:459–473, 1989.

[40] N. Japkowicz, S. J. Hanson, and M. A. Gluck. Nonlinear autoassociation is not equivalent to PCA. *Neural Computation*, 12(3):459–473, March 2000.

[41] A. R. G. Castro, V. Miranda, and S. Lima. Transformer fault diagnosis based on autoassociative neural networks. *2011 16th International Conference on Intelligent System Application to Power Systems (ISAP)*, pages 1–5, 2011.

[42] J. Krstulovic, V. Miranda, A. J. A. Simoes Costa, and J. Pereira. Towards an Auto-Associative Topology State Estimator. *IEEE Transactions on Power Systems*, (99):1–8, 2013.

[43] D. Demuth and M. Beale. *Neural Network Toolbox For Use with MATLAB*. The MathWorks, 4.0.4 edition, 2004.

Appendix A

Literature Review on Distance Protection

Table A.1: Fault Classification.

Paper	Method and Inputs	Test System	Data Sample	Results
[12]	A fuzzy logic method is implemented considering the line current signals as inputs.	The system in study is a transmission line with double end in-feed.	The data sample is obtained using PSCAD\ EMTDC and is composed by 2400 test cases considering different fault types, fault distances, fault inception angles, fault resistances and loading levels.	The method reached a 97% accuracy and a decision time of 10 ms.
[13]	A Fuzzy ART Neural Network (combined use of neural networks and fuzzy logic) is used considering the line current and voltage signals as inputs.	The system in study is a meshed system (CenterPoint Energy STP-SKY).	The data sample is obtained using ATP (Alternate Transient Program) and is composed by 3315 training patterns considering different fault types, fault distances, fault inception angles and fault resistances; 20000 test cases, considering different fault types, fault distances, fault angles and fault resistances, separated in four sets: nominal system - 5000 patterns, weak in-feed - 5000 patterns, off-nominal voltage - 5000 patterns and off-nominal system frequency - 5000 patterns; 4000 patterns are used to tune the parameters of the fuzzy classifier.	In this paper different approaches were tested. The best error achieved in fault type identification was 0%, considering fault type and section classification the error was 1,52%.
[14]	A support vector machine is used for fault classification considering as input the current and voltage signals.	The system in study is a transmission line with double end in-feed.	The data set is composed by 300 cases for training and 200 for testing considering different fault inception angles, fault resistances, source capacities and fault locations.	The method achieved an error less than 3%.

Table A.2: Fault Distance Estimation.

Paper	Method and Inputs	Test System	Data Sample	Results
[15]	A single neuron feed-forward on-line trained neural network is used considering as input an adaptive data window, composed by digital signals, obtained from the derived differential equation line model.	The system in study is a transmission line with double end in-feed.	The data sample is obtained using PSCAD\ EMTDC and considers different fault types, remote- end in-feeds, different fault locations, fault resistances and load conditions.	The method achieves an accurate fault distance estimation within one cycle of the fundamental frequency after the detection of the fault.
[16]	A feed-forward neural network is used considering as input the current level in each phase.	The system in study is a transmission line with double end in-feed.	The data sample only considers open conductor faults at different fault locations and pre-fault current loading.	The method achieves an error of 7% in the distance estimation.
[17]	A fuzzy logic method is used considering as input the line currents and voltages.	The system in study is a series compensated transmission line with double end in-feed.	The data sample is obtained using PSCAD\ EMTDC and considers different type of faults at different locations.	The method has a maximum relative error of 10% and takes 25-50 ms to estimate the fault distance.

Table A.3: Fault Location.

Paper	Method and Inputs	Test System	Data Sample
[18]	A fuzzy neural network model is used considering the fundamental and DC components of the current and voltage signals extracted using a Kalman filter as inputs.	The system in study is a transmission line with double end in-feed.	The data sample is obtained using PSCAD\ EMTDC and considers data from different fault types, fault locations, fault inception angles, fault resistances and loading conditions. The fault classifier training set has 49 samples and each distance estimator has a training set of 56 samples.
[19]	A radial basis function neural network model is used considering the fundamental and DC components of the current and voltage signals extracted using a Kalman filter as inputs.	The system in study is a transmission line with double end in-feed.	The data sample is obtained using PSCAD\ EMTDC and considers data from different fault types, fault locations, fault inception angles, fault resistances, source impedances and loading conditions. The fault classifier training set has 40 samples and each distance estimator has a training set of 70 samples.
[20]	The mathematical morphology method, euclidean norm and the differential equation of the circuit model are used. The input is composed by data windows of phase currents and three line-to-neutral voltages.	The system in study is a transmission line with double end in-feed.	The data sample is obtained using PSCAD\ EMTDC and considers different type of faults at different locations.
[21]	High order statistics (HOS) and neural networks are used considering only the three phase voltage signals as inputs.	The system in study is a transmission line with double end in-feed.	The data sample is obtained using MATLAB Simulink and SimPowerSystemToolbox and considers data from different fault types, fault distances, fault resistances and fault inception angles.

Table A.4: Adaptive Zone.

Paper	Method and Inputs	Test System	Data Sample
[22]	Neural networks are used to calculate the boundaries of the trip region. The inputs during training vary from boundary to boundary and include the active and reactive power flow, different fault resistances and fault reactance. During normal operation the trip decision is made considering as input the apparent impedance measured and the active and reactive pre-fault power flow.	The system in study is a meshed three source system.	The training and test samples are obtained using PSCAD\ EMTDC and consider only phase-ground faults with different fault distances in the line being protected, different fault resistances and pre-fault loading conditions.
[23]	The trip boundaries are defined by mathematical laws based on the identification of the ideal trip regions for typical power system conditions. The input is composed by currents and voltages at the relay location and the equivalent impedance at the wind farm bus.	The system in study is transmission line with double end in-feed composed by two sources: a wind farm and an equivalent system generator.	The data sample considers only phase-ground faults with varying wind farm loading levels, voltage levels, source impedances and system frequencies.
[24]	Radial basis neural networks are used to calculate the boundaries of the trip region. The input signals are the voltage and current signals at the relay location and pre-fault active and reactive power flow.	Two systems are studied: parallel transmission line with double end in-feed and meshed three source system.	The data sample is obtained using PSCAD\ EMTDC and considers only phase-ground faults in different locations with different fault resistances, fault inception angles and system operating conditions. The mutual coupling effect is taken into consideration.
[25]	Neural networks are used to define the trip boundaries. The input consists of pre-fault active and reactive power flows used in combination with fault resistance or fault reactance to define the different boundaries.	The system in study is parallel transmission line with double end in-feed.	The data sample is obtained using MATLAB Simulink and SimPowerSystem-Toolbox and considers only phase-ground faults faults with different locations, fault resistances, fault inception angles and varying system operating conditions. The mutual coupling effect is taken into consideration.

Table A.5: Zone 3 Unintended Tripping.

Paper	Method and Inputs	Test System	Data Sample
[26]	An adaptive algorithm was proposed that uses the derivative of the voltage and the relation between the apparent impedance and the zone of protection to discriminate between normal operation and voltage instability cases.	Two systems are studied: 15-bus system developed by the authors and the Nordic32 system.	The simulations were executed in SIMPOW.
[27]	The method proposed is based on combining the steady-state components with the transient components using a state diagram to discriminate intended trip situations from heavy loading, voltage and transient instability.	The system in study is a parallel transmission line with double end in-feed.	The simulations were executed in PSCAD\ EMTDC.
[28]	The method proposed is based on the combination of the line outage distribution factor (LODF), generation shift factor (GSF)-based power flow estimation method and a secure peer-to-peer (P2P) communication to secure time to perform remedial control actions by a defence system during cascaded events.	The system in study is a six-bus system.	The simulations were executed in MATLAB Simulink and SimPowerSystemToolbox.
[29]	A synchrophasor state estimator taking as input the phasor measurements from different bus locations in the system to discriminate fault conditions from heavy loading, voltage instability and power swing cases.	The system in study is a ten-bus and four generators system.	The simulations were executed in an EMTDC program.

Table A.6: Other Applications.

Paper	Method and Inputs	Test System	Data Sample
[30]	A wavelet transform considering the current and voltage signals at the relay location as input to detect faults by issuing a binary output signal.	The system in study is a transmission line with double end in-feed.	The data sample is obtained using PSCAD\ EMTDC and considers solid ground faults, phase faults, high impedance ground faults, non-linear ground faults and different loading levels.
[31]	A wavelet transform considering the currents and voltages at the relay location is used discriminate power swings from faults.	The system in study is a parallel transmission line with double end in-feed.	The data sample is obtained using PSCAD\ EMTDC and considers different swing conditions, different slip frequencies, different types of fault and fault resistances occurring at different locations.
[32]	An daptive neurofuzzy inference system (ANFIS) is used to detect faults. The input is composed by the trajectory of the measured impedance and fault currents amplitude.	The system in study is a meshed system with 3 generators and 4 lines.	The data sample is obtained using PSCAD\ EMTDC and considers different types of fault, fault locations (inside the protected zone and outside the protected zone), fault resistances, fault inceptions angles and system impedance ratios.

Appendix B

Test System Characteristics

System Power Base = 100 MVA.

Table B.1: Generators.

Name	Bus	Pmax (MW)	Pmin (MW)	Qmax (Mvar)	Qmin(Mvar)	X”d (pu)
G1	1	120	30	60	-60	0,2
G2	2	120	30	60	-60	0,2
G3	3	120	30	60	-60	0,2

Table B.2: Lines.

Name	From Bus	To Bus	Sn (MVA)	Un (kV)	R(Ω/km)	X (Ω/km)	Ysh (mS/km)	Length (km)
Line 1	1	4	100	220	0,2484	0,7888	0,0029	150
Line 2	2	5	100	220	0,2484	0,7888	0,0029	100
Line 3	3	6	100	220	0,2484	0,7888	0,0029	100

Table B.3: Transformers.

Name	From Bus	To Bus	Sn (MVA)	Usec/Upri (kV)	X(%)
T1	1	4	200	150/220	10
T2	2	5	200	150/220	10
T3	3	6	200	150/220	10

Appendix C

Mho Relay Setting

The setting of the mho relay is in accordance with the stated in 2.3.2 and 2.3.3. Since only 3-phase faults are considered in this study the RCA is equal to

- Zone 1: this zone is intended to protect the first 85% of line 1, to do so it was determined through a process of trial and error that the correct reach is 0.869 of the impedance of line 1:

$$Z1_{reach} = |0.869 * Z_{line1}| = 107.79780\ (\Omega) \tag{C.1}$$

$$Z1_{angle} = \angle(0.869 * Z_{line1}) = 72.52033(^{\circ}) \tag{C.2}$$

- Zone 2: 120% of line 1:

$$Z1_{reach} = |1.2 * Z_{line1}| = 148.85771\ (\Omega) \tag{C.3}$$

$$Z1_{angle} = \angle(1.2 * Z_{line1}) = 72.52033(^{\circ}) \tag{C.4}$$

- Zone 3: according to 2.3.2 this zone should be setted to protect the whole length of line 3 considering the worst in-feed conditions, in this case this approach would lead to an enormous dimension of zone 3 due to the variety of in-feed conditions simulated. To establish a comparison basis for the performance analysis of the mho distance relay and neural network based fault locator, the in-feed considered in zone 3 is such that the number of incorrect zone 3 detections (heavy loadings and faults in line 2 detected in zone 3) is equal in the two cases:

$$In - feed = 16.68 \tag{C.5}$$

$$Z1_{reach} = |Z_{line1} + In - feed * Z_{line3}| = 1503.4629\ (\Omega) \tag{C.6}$$

$$Z1_{angle} = \angle(Z_{line1} + In - feed * Z_{line3}) = 72.52033(^{\circ}) \tag{C.7}$$

Appendix D

Data Set

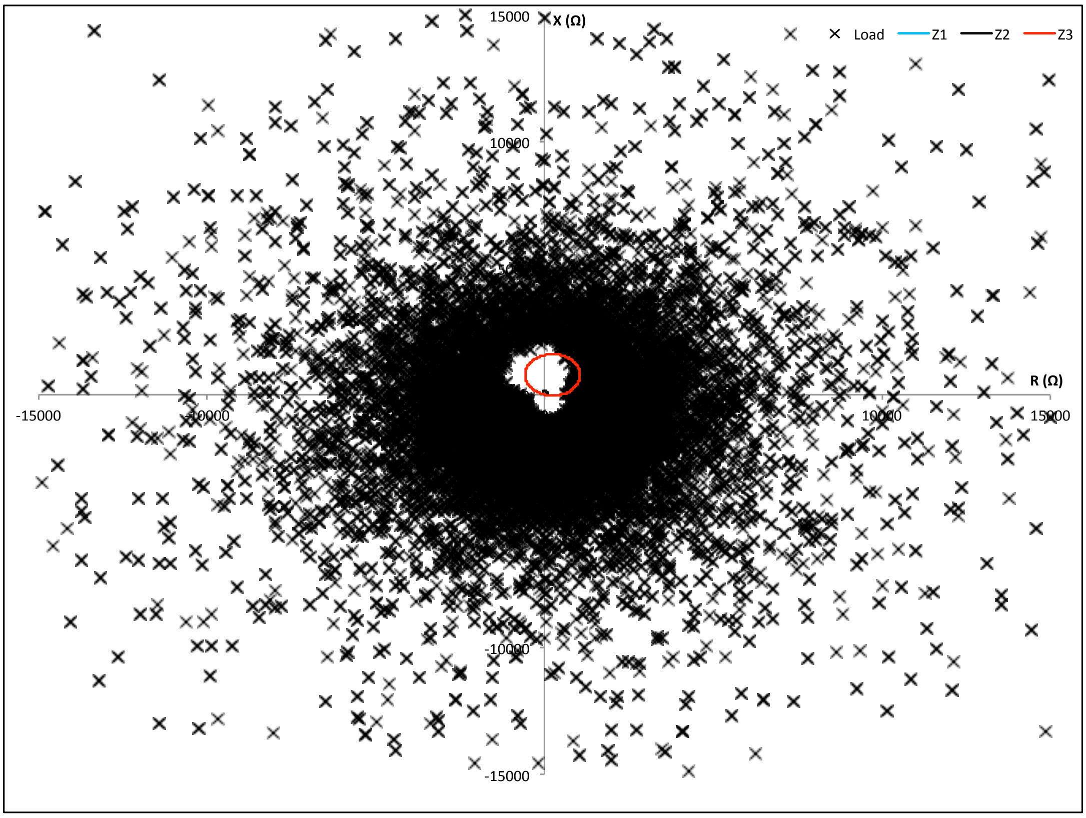

Figure D.1: Power Flow Impedance Points.

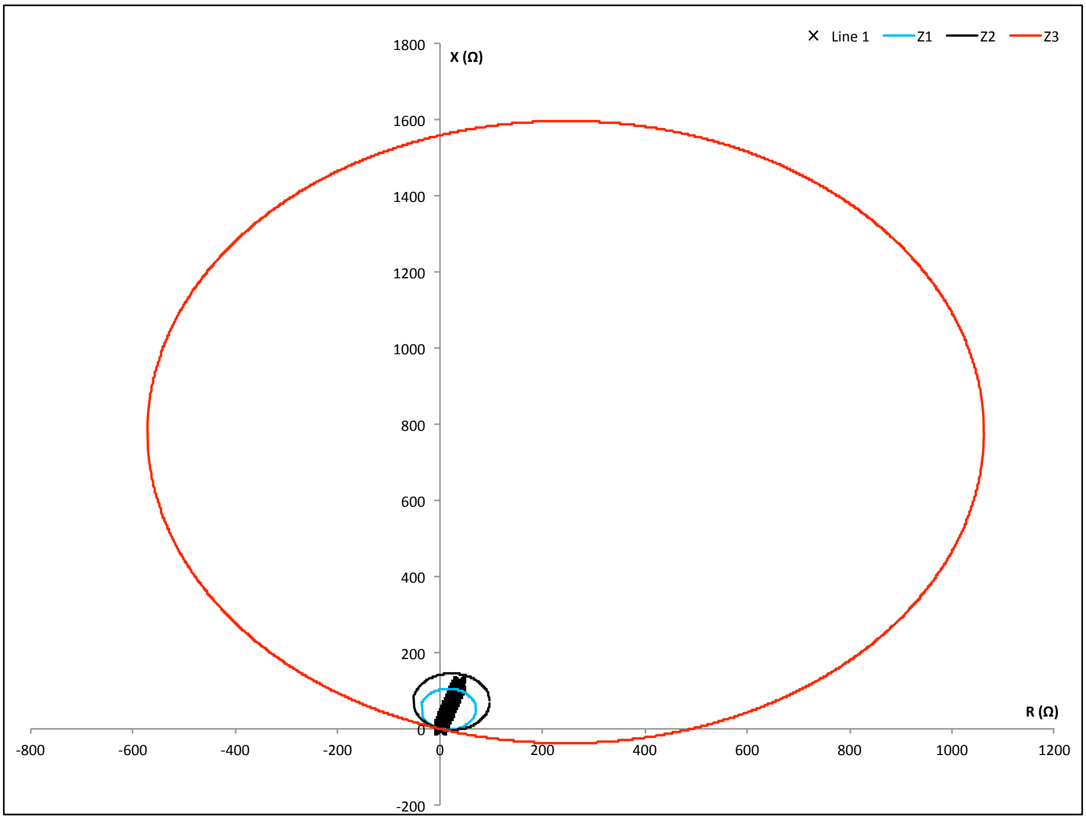

Figure D.2: Line 1 Impedance Points.

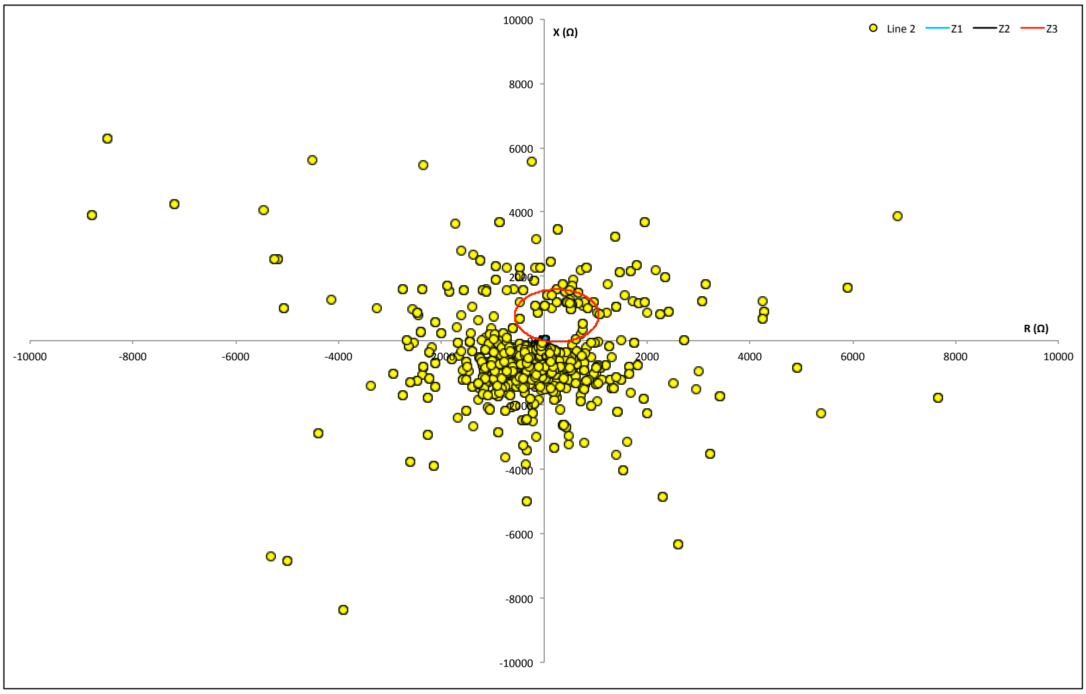

Figure D.3: Line 2 Impedance Points.

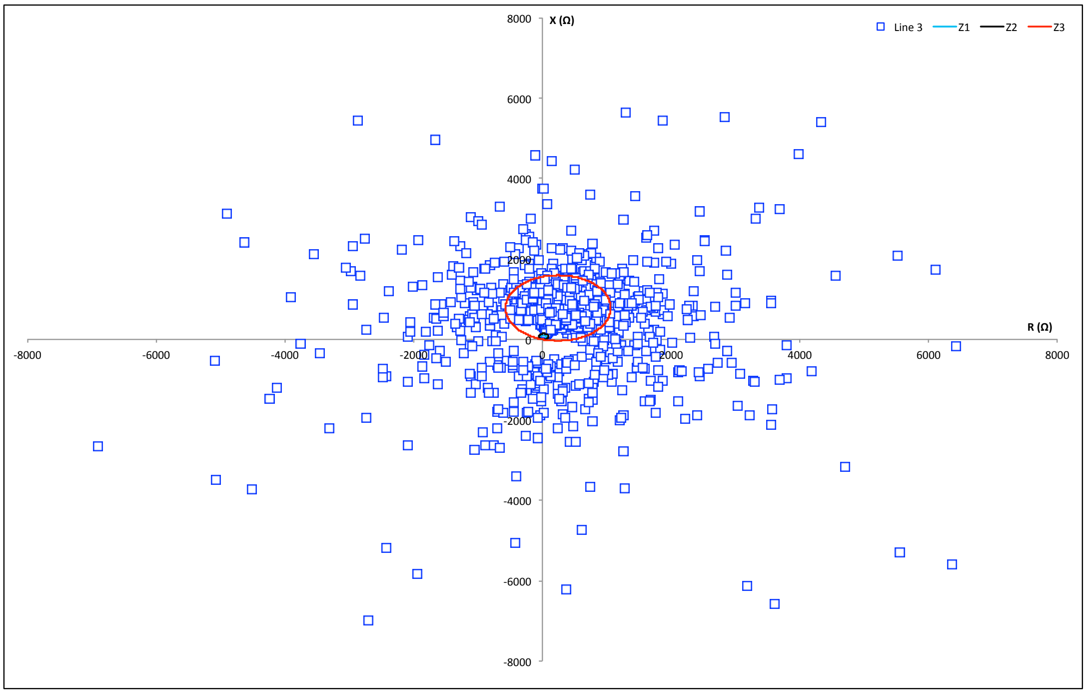

Figure D.4: Line 3 Impedance Points.

Neural Networks Improving the Performance of the Distance Protection

Luis Barreira, Vladimiro Miranda and Hélder Leite

***Abstract*—Neural networks have proven to be an efficient method for classification. This characteristic makes them adequate to be applied in determining the correct operation of distance protection relays since it is basically a classification procedure.**

In this paper, in form of *long abstract* , the efficiency of neural networks in improving the distance protection performance is evaluated. To do so, a data set is created encompassing different pre-fault and fault conditions for a given test system with a distance relay installed in one of the lines. The data set created is used for training a neural network based fault locator that identifies the location of the fault and determines the correct operation of the distance protection. To finalize the performance of the neural network scheme developed is compared with the performance of an mho relay.

The results show that neural networks are an efficient tool for improving the performance of the distance protection.

***Index Terms*—Neural Networks, Pre-fault Load Condition, Intermediate In-feed, Heavy Load**

I. INTRODUCTION

COMPUTATIONAL intelligence techniques have been largely applied in solving power system problems due to their unique capacities. Neural networks are one of these techniques. Between their characteristics, the capacity for performing input-output mapping and recognizing non-linear patterns make them suitable to solve the problem of determining the correct operation of distance protection [1].

Distance protection, or distance relay, is the basis of transmission lines protection. This type of protection system is responsible for detecting faults in power system lines and adopting the necessary actions to isolate the faulted line as quickly as possible. The correct operation of this device is critical to guarantee the security and reliability of the power system.

The problem of determining the correct operation of the distance relay involves classifying the zone where the fault occurs. This procedure is influenced by different internal and external factors that hamper the process of classification and cause the incorrect operation of the distance relay. The relay operates incorrectly when the apparent impedance seen by the relay is different from the real short-circuit impedance between the relay and fault locations, these situations can be classified into two groups[2]:

- **Underreach**: the relay does not operate for a disturbance inside its operation zone. The relay does not operate when it should.
- **Overreach**: the relay operates for disturbances external to its protection zone. The relay operates when it should not.

In this paper the influence of external factors, namely pre-fault load flow, intermediate in-feed and heavy load, is studied and a neural network based fault location scheme is proposed to improve the performance of the distance protection. To create and evaluate the performance of the solution proposed a data set is generated encompassing different pre-fault and fault conditions for a given test system with a distance relay installed in one of the lines. The data set created is then used for training the neural network scheme. To finalize the performance of the solution proposed is compared with the performance of a mho relay.

II. FAULT SIMULATION MODEL

The focus of this dissertation is to analyse the effects of external conditions, namely, the pre-fault conditions, in-feed and heavy load, on the relay performance and propose an optimized distance protection scheme. For this analysis a data set was created based on the following assumptions:

- The data sample must cover a wide range of possible pre-fault loading conditions and line fault locations.
- The system topology must include parallel lines and intermediate in-feeds.

To obtain a data set according to the factors indicated above, a simulation was created, inspired on the monte carlo method, which consists of randomly sampling different operating points and line faults of a given test system. The test system is a meshed system, composed by three generators, three lines, three power transformers and three loads. The relay dimensioned has an mho characteristic and is located in line 1 (see figure 1). Its important to refer that only 3-phase line faults were included in this study due to time limitations. The simulation model was implemented in Matlab.

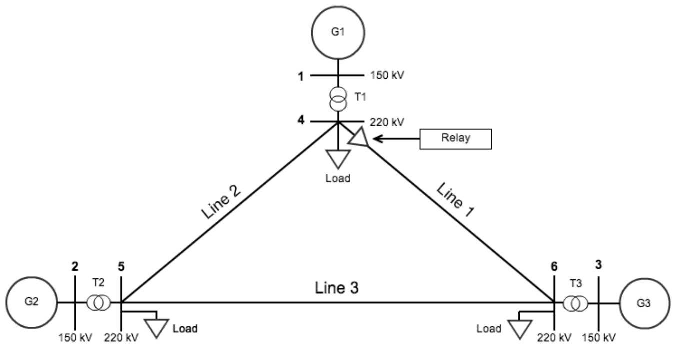

Figure 1: Test System.

ee080202@fe.up.pt
vmiranda@fe.up.pt
hleite@fe.up.pt

The algorithm behind the sampling process consists in the following steps:

Operating Condition Sampling
This step consist of determining a system operating condition by randomly defining the loading level, the power consumed by each load, the active power generated by each generator and the specified voltage in each source:

- Loading level: the total active and reactive power consumed by the loads is defined by randomly sampling a number between the system maximum and minimum generating capacity.
- Load allocation: for each load a randomly sampled percentage of the consumed active and reactive power is assigned. The total power allocated is equal to the system loading level.
- Generation allocation: for each source, the active power produced is randomly defined taking into account that this value must be within the maximum and minimum generating capacity of the generator and the total active power generated must be equal to the total active power consumed.
- Specified voltage: for each generator a specified voltage is sampled. This value varies between 0.9 and 1.1 pu.

Operating Condition Calculation

In this step the operating condition sampled in 1 is calculated using the Newton-Raphson method and the impedance seen by the distance relay is determined.

Operational Limits Verification

To accept the operating condition sampled as a possible hypotheses the following operational limits must be verified, otherwise his power flow is discarded and another one is sampled:

- Generator reactive power production limits (the reactive power generated by each generator is a result of the power flow, so it is not limited in the sampling process).
- Slack bus generator active power production limits (this value is a result of the power flow and depends on the system losses so it is not limited in the sampling process and may exceed the maximum production limit of the generator).
- Voltage limits at the load buses (the voltage limits considered for the load buses are 0.85 pu and 1.15 pu).
- Lines capacity (the power flowing through the lines must be within their capacity limit).
- Transformer capacity (the power flowing through the transformers must be within their capacity limit).

Fault Condition

The last step includes sampling the location of the 3-phase fault and determining the impedance seen by the relay. The fault sampling process includes determining the line and distance in which the fault occurs. The distance is a value that defines the percentage of line between the bus with lower number connected to the line in question and the fault location. With this approach all possible relay tripping decisions are sampled.

III. Towards a Neural Network Fault Locator

Fault location algorithms (see [3]) are used to determine the location of the fault, so that the service restoration time and reliability of the system are improved. In this dissertation a scheme inspired in fault location algorithms is applied to improve the operation of the relay. The basic concept includes detecting a fault, identifying the line faulted, determining the location of the fault within the line and determining the zone.

The effects of intermediate in-feeds and pre-fault loading condition cause the clusters of faults from the different lines, when represented in an R-X diagram (where the operating characteristic of the relay is defined), to define overlapping zones with non-linear boundaries and disjoint. To obtain a better classification performance the solution proposed must be able to define non-linear boundaries and adapt itself to the pre-fault operating condition. In respect to the distance estimation process the solution must be able of determining the location of the fault in the line, from the relationship between the measured voltage and current signals at the relay location. The determination of this value for faults in line 1 is a simple procedure and possible to be executed with a simple mathematical demonstration, for lines 2 and 3 to obtain the distance of the fault in the line the in-feeds from the different sources must be known. Since communication links are not considered, the method used must be able to determine this value with only local information.

To consider the pre-fault condition effect the input consists of a data window of two sequential measures, namely the voltage and current signals and the angle between them, measured at the operating moment and these same values measured in the sample before. With this approach, when a fault occurs the value measured before is related to the pre-fault operating condition and the value measured in the moment is related to the sub-transitory short-circuit condition. During operation these values are continuously injected in a pipeline mode: when a new sample enters the sample before leaves (see figure 2).

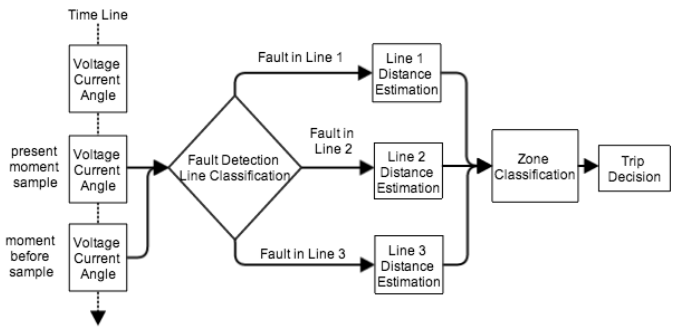

Figure 2: Architecture of the Solution.

The option for neural networks for fault detection, line classification and distance estimation resides in the fact that these method has the aptitude of recognizing non-linear patterns and reproducing non-linear functions that cannot be modelled by mathematical expressions and the capacity of classifying patterns in a more precise way [1].

The fault detection and line classification processes are incorporated in the same neural network. The reason behind the junction of these two steps is that both of them consist of a classification process, besides it was observed through experimentation that this junction does not influence the performance of neither of the steps. For distance estimation a neural network is trained for each of the lines in the system. The neural networks created for both processes are feedforward.

The fault detection and line classification procedure is a classification problem. In feedforward neural networks this type of problems is my be solved by associating a specific codification to each class. During training, different instances belonging to different classes are presented as input, and the desired output is the codification associated with the class of the instance presented. The objective function is the mean square error between the output obtained and the respective target, which is minimized by the training algorithm. The output obtained is not an integer so this value is rounded to obtain a classification.

The feedforward neural network used in the fault detection and line classification block is composed by four layers of neurons, where the first three layers have the same number of neurons and the fourth layer is composed of a single neuron. In the first layer no activation function is applied, in the middle layer the activation function is a hyperbolic tangent and in the output layer the activation function is linear.

The estimation of the distance to the fault is a function approximation problem. Feedforward neural networks, when applied to function approximation, have the objective of reproducing the unknown function behind the relation between the pairs of inputs and outputs of a certain manifold [4]. To learn this relation the minimization of the error between the machine output and the desired response is applied as the objective function.

The feedforward neural network used for the fault distance estimation of each line consists of three layers of neurons where the activation functions for the middle and output layers are: hyperbolic tangent and linear (the input layer has no activation function). For each set of inputs the desired output is the fault distance. The number of neurons in the middle layer was defined by a trial and error process so that a compromise is achieved between the dimension of the hidden layer and the distance estimation accuracy (line 1 has 2 neurons, line 2 and line 3 have both 3 neurons).

IV. Performance Evaluation

In this section the performance of the scheme developed is compared with the performance of a dimensioned mho relay. The Levenberg-Marquardt training algorithm is used in every test. The neural networks were created using the MATLAB neural network toolbox. The data set used for training, test and validation is the one generated in section II. The performance evaluation is performed by analysing the correct responses to a validation set composed by 9939 cases of faults in line 1, 9847 cases of faults in line 2, 10214 cases of faults in line 3 and 30000 cases of normal operation.

The perspective adopted in this evaluation aims at identifying and quantifying the incorrect and correct zone classification situations. The results obtained by each scheme in zone classification are presented in table I where OPZ stands for Out of Protected Zone and NNFL stands for Neural Network Fault Locator:

Table I: Zone classification performances

	Zone Classification		Number of Cases	
	Correct	Detected	Mho Relay	NNFL
Correct Operations	OPZ	OPZ	39756	39756
	Zone 1	Zone 1	8505	8505
	Zone 2	Zone 2	2410	2478
	Zone 3	Zone 3	8176	9041
Incorrect Operations	OPZ	Zone 3	91	91
	Zone 3	OPZ	1062	129

In table I it is observed that the NNFL and the mho relay obtain different classifications in three cases: for faults in Zone 2 classified has Zone 2, for faults in Zone 3 classified has Zone 3 and for faults in Zone 3 classified as OPZ. The first two cases represent correct operations, the difference observed arises due to the assumptions admitted in classifying a certain operation as correct or incorrect for faults in the boundary of zone 2 and 3 and show that for the NNFL Zone 2 has a longer reach. The last case represents incorrect classifications where the fault occurs in line 3 and the relay does not issue a trip decision. In this case the NNFL obtained a better performance, reducing the number of incorrect cases from 1062 to 129.

The cases where the correct classification is OPZ and the classification obtained is Zone 3 also represent incorrect decisions. Although the number of occurrences is equal for both of the schemes, the cases that they represent are different: the mho relay is composed by 39 cases of faults in line 2 and 53 heavy load conditions while the NNFL is composed by 91 cases of faults in line 2.

V. Conclusion

Neural networks performing pattern recognition are efficient in discriminating non fault from fault conditions (100% of correct decisions for the validation set tested), thereby eliminating the cases of incorrect operation due to heavy load conditions.

The solution proposed improves the performance of the relay in zone 3, increasing the number of correctly detected zones in 933 cases when compared to a mho relay for the validation set tested.

References

[1] S. S. Haykin, *Neural Networks - A Comprehensive Foundation*, 2nd ed. Delhi, India: Pearson Prentice Hall, 1999.

[2] S. H. Horowitz and A. G. Phadke, *Power System Relaying*, 3rd ed. John Wiley and Sons, 2008.

[3] P. K. Dash, A. K. Pradhan, and G. Panda, "Application of minimal radial basis function neural network to distance protection," *IEEE Transactions on Power Delivery*, vol. 16, no. 1, pp. 68–74, 2001.

[4] J. C. Principe, N. R. Euliano, and W. C. Lefebvre, *Neural and Adaptive Systems: Fundamentals Through Simulations*, J. Wiley and Sons, Eds., 2000.