Development of an Artificial Neural Network based Hardware Prototype for Fault Localization in Distribution Grids

M Kenan Mehmood, Björn Keune
Institute of Energy Systems, Energy Efficiency and Energy Economics (ie^3)
Technical University Dortmund, Germany
kenan357@yahoo.com, bjoern.keune@tu-dortmund.de

Abstract— **In modern distribution grids, the task of fault localization using conventional techniques is increasingly becoming a challenge due to the rising domination of inverter-based, volatile, distributed power generation. Since approved methods from high voltage level such as reactance-based methods lack accuracy in distribution level topology, alternative approaches for accurate fault localization are required. Within the scope of this work, an artificial neural network (ANN) based solution for the localization of electric faults at distribution level has been developed, evaluated and implemented on standard hardware from industrial automation technology i.e. a programmable logic controller (PLC). A reduced yet representative model of a distribution grid incorporating a variety of aspects influencing the accuracy of fault localization such as distributed generation, ring network topology with open or closed loop as well as variable fault resistance has been developed. Current and voltage measurements generated under various fault conditions have been used for training of an ANN. Different ANNs have been trained with various network structures and training algorithms and after thorough analysis and comparison of their performance, the most suitable networks have been implemented on hardware and tested in hardware-in-the-loop configuration. Thereby a real-time simulator suitable for application testing and rapid prototyping provided process values of the modeled distribution grid.**

Keywords: Distributed Generation, Fault Localization, Artificial Neural Network, Programmable Logic Controller

I. INTRODUCTION

During the last decade there has been a development towards building smaller generation units which are connected directly to distribution networks, designated as distributed generation (DG) [1]. Generally, DG units with a generation capacity of up to 10MW are connected directly to the distribution network [2]. However, adding extra generators in a power system designed for a fixed number of sources raises the system short circuit level, on the basis of which the whole power distribution switchgear is designed.

Another problem arising from having DG in a distribution network is the detection and localization of faults in the system. In case of a fault, the DG unit connected to a feeder feeds a fault on another feeder thus resulting in tripping of the healthy feeder as well [3]. The decision to operate a network in open or close loop further complicates the situation as fault current has even more paths available to flow from DG towards the location of fault. Reference [4] provides a comprehensive explanation of the impact a DG source can have on a distribution network. Not only does it compromise the very integrity of distribution switchgear in case of short circuit but may also result in maloperation of protective relaying schemes and tripping of multiple feeders. This can make the job of fault localization all the more challenging and calls for innovative solutions in order to ensure quick power restoration after an interruption.

The purpose of this work is to design and evaluate a hardware prototype for precise fault localization in modern distribution networks using techniques from the field of computational intelligence. References [5], [6] and many other similar works have proven the practicability of fault localization in transmission lines using artificial neural networks (ANNs) with a high degree of accuracy. However, compared to a transmission line, a distribution network consists of more feeding circuits and has a non-unidirectional flow of current in case of a fault.

The presented work is intended to explore practicability and performance of ANNs for fault localization in a distribution grid and to implement them in hardware which can be used in conjunction with the existing switchgear in a substation. Thereby a distribution grid model is simulated for various fault cases and the generated voltage and current values are used for training of an ANN. The trained network is then implemented on hardware and its performance is evaluated by means of a hardware-in-the-loop (HiL) test whereby the developed hardware is connected with a real-time simulator running the grid model. The purpose of HiL testing is to evaluate the feasibility of the said prototype to serve as a standalone device which can be networked with distribution switchgear.

II. DISTRIBUTION GRID MODEL

Fig. 1 shows the elementary single line diagram (SLD) of the distribution network to be modeled. Only measurement devices (current and voltage transformers) relevant for fault localization have been depicted in the SLD. No protection devices and protection schemes have been shown since system protection is not the primary focus of this work. A Simulink® model based on this SLD has been developed.

Since faults are to be simulated only on the distribution line, the loads being fed have been modeled as lumped loads. The coupling circuit breaker allows forming a closed loop with the other distribution circuit. DG source gets a tie in to the distribution line 1 where first load transformer is located.

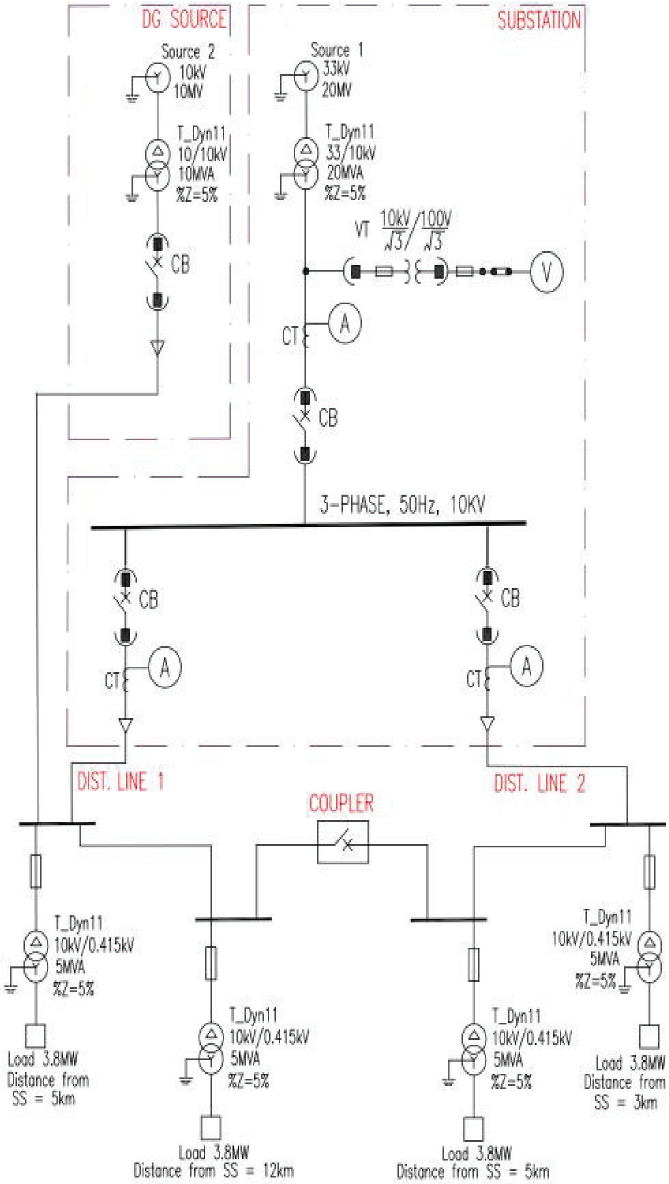

Figure 1. Distribution system single line diagram

In Fig. 1, switchgear feeding the distribution grid and the upstream network has been modeled by a 33kV generator connected to a 33/10kV step down transformer. A combined heat and power (CHP) generation facility has been modeled by a 10 kV, 10 kVA generator, designated as 'DG Source'. A 1:1 transformer with a delta connected winding on the generator side is connected to prevent ground faults from travelling to the generator side. 'Dist. Line1' and 'Dist. Line2' are underground cable distribution system lines modeled by three phase pi-section line masks. In order to generate the training data for ANN, the model has to be simulated multiple times with different simulation parameters. Each simulation has to be run for 0.4 seconds i.e. 20 AC cycles. In every simulation, fault is introduced after 10 cycles in phasor simulation mode. Current and voltage samples generated during this time are enough for ANN training.

III. Grid Model Simulation and Fault Data Generation

After developing a comprehensive model incorporating aspects like distributed generation and closed loop operation which affect the accuracy of conventional fault localization methods, it has to be simulated repeatedly with various fault locations and fault resistances. The resulting current and voltage measurements are to be stored and later used for development of a mathematical model for fault localization by ANN training.

A. Model simulation

Following aspects have been taken into consideration for simulation of grid model in order to generate fault data for ANN training.

Table 1. Model parameters for fault simulation

Simulated faults	Phase to ground, Phase to phase, Double phase to ground, Three phase
Range of fault resistance	0.01Ω, 0.05Ω, 0.1Ω, 0.5Ω, 1Ω, 2Ω, 5Ω, 10Ω, 15Ω, 20Ω, 25Ω, 30Ω, ……, 75Ω
Voltage variation	In addition to nominal system voltage (10kV) -5%, -2.5%, -1%, +1% and +2.5%
Fault locations	An incremental distance of 0.2 km (5 faults locations per km of line) along the complete length of the distribution lines
Open loop operation (coupling circuit breaker open)	Fault measurements obtained by running a separate simulation for each feeder. A distinct neural network is to be trained for each feeder since both the feeders generate completely different fault measurements
Close loop operation (coupling circuit breaker closed)	Both distribution lines joined together to form a ring. Both feeder currents and voltages are to be obtained by running a single simulation with feeder 1 as the starting point and feeder 2 as the end point of the loop
Fault data recording location	Outgoing feeder currents and voltages in substation. Hardware is to be installed in the substation and receives the required data from relays for fault location estimation thorough serial communication / networking

B. Fault data generation

Following is a brief description of the measurements needed and how they are organized in a single input data matrix for ANN training.

Table 2. No. of model simulations for each fault type

Operation	Dist. line	Measurement points / km (A)	No. of voltage levels (B)	No. of fault resistance values (C)	Length of dist. line (km) (D)	Total no. of simulations to be run (A x B x C x D)
Open loop	1	5	6	21	12	7560
	2	5	6	21	8	5040
Closed loop		5	6	21	20 (12 + 8)	12600

- Feeder current and voltage has 3 phases namely $\mathbf{I_a}$, $\mathbf{I_b}$, $\mathbf{I_c}$ and $\mathbf{V_a}$, $\mathbf{V_b}$ and $\mathbf{V_c}$ which can be stored in **6** variables.
- Current and voltage phasors have real as well as imaginary parts. With voltages and currents of all phases having 2 parts, the input vector now has **12** columns.
- Both pre- and post-fault values of all currents and voltages are to be recorded. Therefore two variables (for pre fault and post fault values) have to be defined for both real and imaginary parts of every phase of current and voltage vectors. This increases the dimensions of measurement matrix to **24** distinct columns.

Table 2 provides an exact count of the number of times the grid model has to be simulated in various configurations with various combinations of model parameter values. The final count in each case gives the number of rows that will be added to the measurement matrix which has 24 columns as described above.

IV. Development of a Mathematical Model for Fault Localization by ANN Training

The objective in this step is to use the generated fault data to establish a mathematical relation between the observed effects (i.e. current and voltage measurements generated as a result of a fault) and causes of these particular effects (i.e. fault locations). A fitting function to model the system behavior has to be found using ANN training. Once the system is identified and a fitting function is developed, it can be used for determining fault locations (causes) using fault measurements (effects) that are not part of the previous knowledge base.

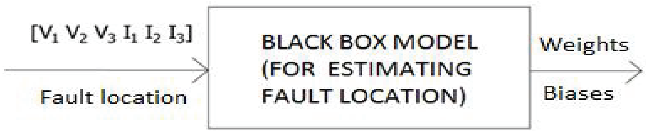

Figure 2. Graphical representation of ANN training

A. *Quantification of training tasks*

Since, in addition to fault location, the fault data also contains information about the fault resistance, it is worthwhile to train another ANN that provides a fit between fault measurements (24 inputs) and both fault location and fault resistance (2 outputs). Fault resistance data can be used for fault analysis and be included in public disturbance records. Valuable historic data like this can also prove helpful in later research. Since ANNs with both one output (fault distance only) and two outputs (fault distance as well as resistance) are to be trained for **12** cases (4 faults for 3 feeder configurations) the total number of scenarios becomes **24**. For every test case, two ANNs are to be trained (using the same training data) and their outputs are to be averaged. Doing so improves the overall accuracy of ANN model output. Since two ANNs with their output averaged are to be trained for each scenario, the final count of ANNs to be trained stands at **48.**

B. *ANN training parameters*

Before creating a feedforward ANN object, its structure has to be specified. The default feedforward net object has 2 layers. 70% of the training data has been allocated for training whereas 15% each is used for validation and testing. Division of data is random. Instead of selecting random initial values for weights and biases for neurons in the net, Nguyen-Widrow initial conditions have been used for both layers of the net thereby reducing the training time by an order of magnitude [7].

Owing to good convergence and fitting characteristics in all the trials, a neural net with 24 inputs, 15 inner layer neurons, and (depending on the number of outputs) 1 or 2 neurons in the output layer is to be used for training of ANNs for fault location estimation. Levenberg-Marquardt optimization with bayesian regularization is the training algorithm of choice [8].

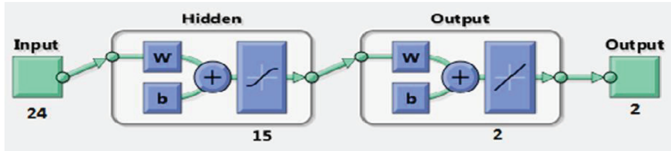

Figure 3. Graphical representation of ANN training

C. *ANN training results*

Having finalized all the training parameters, network structure and training algorithm, development of mathematical models for fault localization is then carried out. Target value of mean squared error (mse) is 0.0005. In some cases, training has been allowed to continue beyond this error value as long as no over fitting is observed. ANNs for both single model output (fault location only) and two model outputs (fault location as well as resistance) have been trained. Training results for all the faults have been presented in the Table 3.

For A-B fault, the satisfactory training performance can only be achieved for low resistances (up to 5Ω) due to the fact that for higher resistance values, the voltages of phase A and B reach close to that of phase C (the healthy unfaulted phase). Since all voltage values become almost constant with very little variation and accordingly very low fault current flows (almost close to normal load current), it becomes difficult to find a fitting function for a system which, apart from some large voltage and current variations for low value of a parameter (resistance), is almost constant.

This difficulty can be avoided by considering the fact that in an underground cable distribution system, usually a fault between two phases occurs when the insulation between them breaks down which leads to a low resistance fault. An arcing fault can be encountered when the insulation breaks down only partially. Reference [9] provides an estimate of arc resistance in overhead distribution system (1.1Ω) which usually depends on many other parameters and in overhead systems, the separation between two phase conductors is much larger as compared to two cores of a cable. This observation therefore justifies limiting the upper bound on the fault resistance value for A-B fault to 5Ω which is still way above the arc resistance value. Since fault resistance has to be limited to only 5Ω for A-B faults, the grid model has to be simulated for more values of fault resistance between 1Ω and 5Ω. The additional values in this range for which faults on the distribution line are introduced are 1.5Ω, 2.5Ω, 3Ω, 3.5Ω, 4Ω, 4.5Ω in order to ensure better fitting and generalization.

Table 3. ANN training errors for all faults in open and close loop

MSEs of ANN Training for A-Gnd Fault Localization		
Configuration and feeder number	Output: Fault location only	Outputs: Fault resistance and location
Open Loop - feeder 1 (ANN # 1)	0.000248	0.000313
Open Loop - feeder 1 (ANN # 2)	0.000344	0.000366
Open Loop - feeder 2 (ANN # 1)	0.000099	0.000101
Open Loop - feeder 2 (ANN # 2)	0.000099	0.000103
Close Loop (ANN # 1)	0.000099	0.000892
Close Loop (ANN # 1)	0.000099	0.000566
MSEs of ANN Training for A-B Fault Localization		
Configuration and feeder number	Output: Fault location only	Outputs: Fault resistance and location
Open Loop - feeder 1 (ANN # 1)	0.028451	0.012789
Open Loop - feeder 1 (ANN # 2)	0.019033	0.013188
Open Loop - feeder 2 (ANN # 1)	0.000178	0.000089
Open Loop - feeder 2 (ANN # 2)	0.000143	0.000096
Close Loop (ANN # 1)	0.000010	0.000010
Close Loop (ANN # 1)	0.000157	0.000170
MSEs of ANN Training for A-B-Gnd Fault Localization		
Configuration and feeder number	Output: Fault location only	Outputs: Fault resist & location
Open Loop - feeder 1 (ANN # 1)	9.957 x 10^{-9}	0.000139
Open Loop - feeder 1 (ANN # 2)	9.901 x 10^{-9}	0.000141
Open Loop - feeder 2 (ANN # 1)	9.992 x 10^{-9}	0.000099
Open Loop - feeder 2 (ANN # 2)	9.978 x 10^{-9}	0.000094
Close Loop (ANN # 1)	9.992 x 10^{-7}	0.000099
Close Loop (ANN # 1)	9.981 x 10^{-7}	0.000198
MSEs of ANN Training for A-B-C-Gnd Fault Localization		
Configuration and feeder number	Output: Fault location only	Outputs: Fault resist. & location
Open Loop - feeder 1 (ANN # 1)	9.714 x 10^{-9}	--
Open Loop - feeder 1 (ANN # 2)	9.714 x 10^{-9}	--
Open Loop - feeder 2 (ANN # 1)	9.576 x 10^{-9}	--
Open Loop - feeder 2 (ANN # 2)	9.982 x 10^{-9}	--
Close Loop (ANN # 1)	9.943 x 10^{-9}	--
Close Loop (ANN # 1)	9.955 x 10^{-8}	--

As highlighted in Table 3, the target error value (0.0005) has not been obtained when an ANN for faults on distribution line 1 in open loop operation is trained, even though training has been performed for a much lower range of resistance. The reason for this is that the distribution line 1 has a DG source connected to it so the voltage drop in case of a fault with resistance on this particular line is not very high. Only for very low fault resistance is there a real fault condition. The lesser drop in voltage on this line makes it difficult to reduce the error and the training stops after maximum value of 'Mu' (training iteration step) is reached.

Three phase fault is a symmetric fault and when a network for only fault location has to be trained, very good results are obtained. However ANN training for two outputs is not possible. The training algorithm fails to converge. The reason for this is that for all the different values of fault resistance in the ground path, all the phase voltages and currents have same values and fault measurements and change only with either pre fault voltage or fault location. Therefore, for this periodic function, ANN training is very good when a function is to be found that provides a relation between fault measurements and fault location only. However the training fails when the fitting function has to include fault resistance as well since fault resistance has no effect on fault measurements. Thus a relation cannot be established between fault measurements and a combination of fault resistance and location.

D. *ANN model testing*

A test dataset consisting of fault measurements that are not part of training data has been used for testing the fault location estimation capability of the developed model. Test inputs generated for all cases are to be applied to the corresponding developed models and the outputs generated by them are then compared to the true test outputs for calculation of error. The next section presents all the test results obtained.

E. *ANN model test results*

Test results in graphical form have been presented for A-Gnd fault in the figure below.

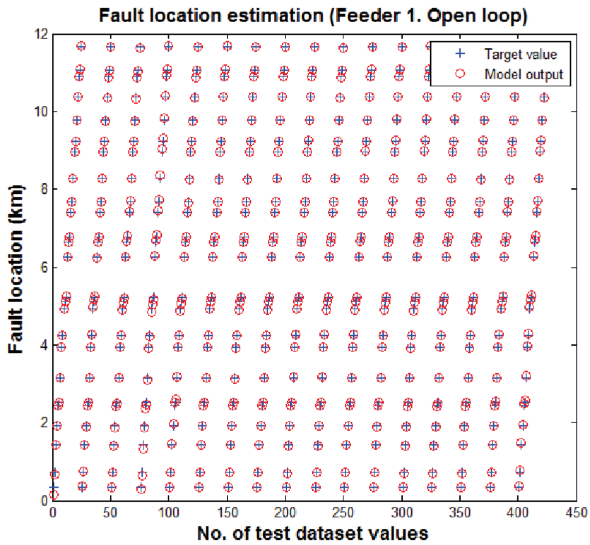

Figure 4. A-Gnd fault locations for various resistances

The plot in Fig. 4 shows the fault locations estimated by ANN model superimposed on the actual fault locations for different values of fault resistance. (17 different values of fault resistance represented by 17 slanted vertical lines formed by fault location points). Results for other faults have been summarized in tabular form. The purpose of presenting fault location plots for at least one of the faults is to provide a vivid illustration of the quality of test results.

Table 4 provides information about the percentage of test cases in all operational scenarios where the estimation error for fault location (and fault resistances) is less than or equal to 5%

Table 4. Summary of test results for all test cases

Oper-ation	Dist. line	Fault type	Total test cases (A)	Cases with > ±5% location error (B)	%age of cases with >95% accuracy $(\frac{A-B}{A})$ %	Cases with > ±5% resistance error (C)	%age of cases with >95% accuracy $(\frac{A-C}{A})$ %
Open loop	1	A-Gnd	425	8	98.12	1 output only (Location)	
	2		289	7	97.58		
Closed loop			714	4	99.44		
Open loop	1		425	6	98.56	21	95.06
	2		289	11	96.19	4	98.62
Closed loop			714	16	97.76	55	92.3
Open loop	1	A-B	275	74	73.09	1 output only (Location)	
	2		187	7	96.26		
Closed loop			462	7	98.48		
Open loop	1		275	67	75.64	79	71.27
	2		187	12	93.58	15	91.98
Closed loop			462	3	99.35	62	86.58
Open loop	1	A-B-Gnd	425	0	100	1 output only (Location)	
	2		289	0	100		
Closed loop			714	0	100		
Open loop	1		425	0	100	12	97.18
	2		289	0	100	4	98.62
Closed loop			714	0	100	11	98.46
Open loop	1	A-B-C-Gnd	425	0	100	1 output only (Location)	
	2		289	0	100		
Closed loop			714	0	100		
Open loop	1		425	Not applicable (Model with 2 outputs can't be trained)			
	2		289				
Closed loop			714				

(in other words, where an accuracy of greater than or equal to 95% has been achieved).

Another observation in Table 4 is that the accuracy of fault location estimation gets better as the degree of symmetry in faults increases. The results for asymmetrical fault (A-Gnd) are also very good. Since the training performance (mse) for A-B fault on distribution line 1 (with a tie-in for DG) in open loop is off by two orders of magnitude from the target mse (0.0005), the percentage of test cases having an error value of less than 5% is also rather low.

The models for which test results have been presented are to be implemented in hardware. These are the models trained using the measurements in substation. Two separate hardware programming projects have been created for these implementations. The implementation with only one output (fault location) is the preferred one for use since it has higher accuracy and works in all scenarios. Furthermore, no model for fault resistance estimation is available for three phase faults.

V. Hardware Implementation of ANN Model and Hardware-in-the-Loop Testing

After development and testing of mathematical models for localization of various faults in a software simulation environment, the final step is to develop a hardware prototype with an ANN model as the underlying algorithm which can be used in conjunction with the existing electrical control gear in a substation. A programmable logic controller (PLC) has been the hardware platform of choice. After the developed model has been implemented on the PLC, its performance has to be tested for various faults and in different modes of grid operation.

In order to test a PLC in a lab, test signals similar to the ones generated by metering devices installed in MV switchgear are required. For this purpose, a real time (RT) simulator simulating the behavior of a real grid is used. The ANN model implemented on PLC is trained using the measurements generated by simulation of a grid model. The same Simulink® model is to be simulated now by the RT simulator for testing of PLC. Data exchange between the hardware under test and the RT simulator takes place via TCP/IP. Doing so eases the limit on the number of values that can be transmitted from the RT simulator to PLC using hardwired analog signals since both of these devices have limited number of analog I/O ports. Moreover, in an analog signal all the values are mapped to a 4-20mA current signal which can result in inaccuracies when a value with a span of 20000V (from -10000V to +10000V) is to be mapped on a signal with a span of only 16mA. In addition to feeder current and voltage values, feeder statuses (fault / no fault) and the state of coupling circuit breaker (open / close loop operation) is also transmitted to PLC.

For ANN model implementation, the available hardware is a Beckhoff C6920 industrial PC (IPC) with the PC based control software TwinCAT 3 for application engineering. The flow chart in Fig. 5 shows the organization of PLC application for ANN based fault localization and the order in which different blocks are called and executed.

First stage deals with establishing TCP/IP communication with the simulator. After a connection is established and the transfer of TCP frames to PLC begins, every incoming frame is processed and required fault data is extracted from the received TCP frame. Along with data extraction and storage in appropriate variables, the PLC keeps on monitoring all the variables in a cyclic fashion.

Whenever a fault is detected, (fault detection variable returns a non-zero value) the fault type and the mode of grid operation are examined and an ANN function with appropriate values of weights and biases is invoked. The number of neurons in the output layer is equal to the number of ANN outputs. PLC applications for both a single output (fault location) and two outputs (fault resistance and location) have been developed. Even though extensive testing of the model performance and accuracy has been carried out in software before hardware implementation could be proceeded with, it is important to test the programmed hardware as well before deployment using some test cases whose results are already known from testing in software. This is to ensure that the translation of Matlab® code of ANN model into IEC 61131-3 structured text (ST) has been carried out correctly. Speed and stability of application can also be put to test this way. Test results show that the fault location estimation error in hardware deviates slightly (1-2%) from the ones obtained in software testing. This is due to the mismatch between the fault measurements generated by Simulink® grid model and the actual grid i.e. model transferred to the RT simulator for HiL testing (considered the real grid).

As far as the speed of computation is concerned, all computations are completed within the default PLC cycle time of 10ms which means that, even though the numerical effort is quite high for matrix operations in ANN model, PLC is a suitable platform for such an application.

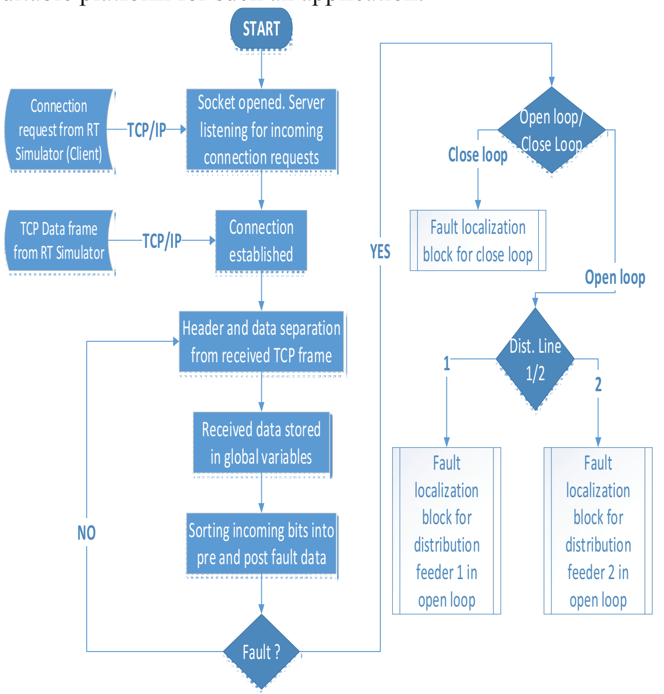

Figure 5. Organization of the PLC application for fault localization

Conclusion

An ANN for fault localization in distribution grids has been designed, implemented on hardware and its characteristics such as accuracy and speed have been evaluated. For training of ANNs, voltage and current measurements in the substation have been used. As hardware platform for the designed algorithm, a PLC has been used. The developed hardware is then connected to the grid model using a real time simulator for HiL testing. The results are encouraging and PLC has proved to be an economical and reliable platform for such applications with real-time capability, reproducibility and deterministic behavior. Using both primary and secondary substation measurements for the training of an ANN can yield even more accurate results. The number of generated fault measurements would then double but this is where TCP/IP communication can come really handy since no additional hardware is required.

In general, the results obtained using ANNs for fault localization have been very encouraging and show that they can be employed as an alternative to the conventional impedance based or travelling waves methods. The greatest effort however is required for ANN training since it has to be performed multiple times when the training algorithm either doesn't converge or gets stuck in a local minimum. The training has to be performed once again if changes are made in the grid structure (e.g. new elements are added). If the size of network increases, the time required for generation of fault measurements and ANN training also increases.

References

[1] Marvik, Jorun I., "Fault Localization in Medium Voltage Distribution Networks with Distributed Generation", Norwegian University of Science and Technology, June 2011J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Dugan, Roger. C. and McDermott, Thomas E., "Distributed Generation." IEEE Industry Applications Magazine, March/April 2002. p. 19-25

[3] Hager, M., Sollerkvist, F. and Bollen. M.H.J. "The impact of distributed energy resources on distribution-system protection." August 10, 2006. NORDAC. Stockholm, Sweden

[4] IEEE – Power system relay committee, "Impact of Distributed Resources on Distribution Relay Protection – a report to the line protection subcommittee of the IEEE power engineering society", August 2004

[5] Chen, Z., Maun, J.C., " Artificial Neural Network Approach to Single-Ended Fault Locator for Transmission Lines", 0-7803-3713-1/97, IEEE 1997

[6] Li, K.K., Lai, L.L., David, A.K, "Application of Artificial Neural Network in Fault Location Technique" 0-7803-5902-X/00, IEEE 2000

[7] Nguyen, D., and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights" Proceedings of the International Joint Conference on Neural Networks,Vol.3, 1990, pp. 21–26

[8] Beale, M.H., Hagan, M.T. and Demuth, H.D., "Neural Network Toolbox™ - User's guide", Matlab® R2014a, Mathworks

[9] Terzija, Vladimir.V, Ciric, R, Nouri.H, "Improved Fault Analysis Method Based on a New Arc Resistance Formula" IEEE transactions on power delivery, Vol. 26, No. 1, January 2011