Real-Time Implementation and Evaluation of a Support Vector Machine Based Fault Detector and Classifier for Distribution Grids

Mohammad Umair Ehsan, M. Sc. Björn Keune
Institute of Energy Systems, Energy Efficiency and Energy Economics (ie3)
Technical University of Dortmund
Dortmund, Germany
mohammad-umair.ehsan@tu-dortmund.de, bjoern.keune@tu-dortmund.de

Abstract—**Bridging the gap between theoretical modeling and practical implementation is essential in fault detection, classification, and location methods for modern distribution grids. Currently distribution grids are characterized by dispersed infeed and distributed power generation that render conventional protection settings more challenging, and, hence, new methods must be investigated to resolve these issues. In this paper, a novel framework capable of detecting and classifying faults in power distribution grids is presented. The proposed algorithm formulates a unique fault classification technique based on measurement samples of three-phase voltage and currents after the occurrence of a fault event in power distribution grid. Thereby, negative sequence components of three-phase voltage and current quantities are used for online fault detection that triggers a fault classification method based on a support vector machine. In order to simulate fault scenarios, a model of a reference distribution grid incorporating distributed generation is developed, simulated, and analyzed under fault conditions. For the final evaluation, the designed protection scheme is implemented on a Programmable Logic Controller connected to the model running on a real-time simulator hardware that provides voltage and current measurements. Thus experimental results are provided to demonstrate and prove the contribution of the algorithm in its ability to correctly identify and classify faults in modern distribution grids.**

Keywords-Distributed Generations (DGs); Secondary Substation (SS); Photovoltaic (PV); Negative Sequence Components (NSCs); Support Vector Machine (SVM); Programmable Logic Controller (PLC); Real Time Simulator (RTS); Hardware in the loop (HIL);

I. INTRODUCTION

One of the most important factors of an electrical power distribution system is its reliability and uninterrupted supply for its end users. Over the past several decades there has been rapid growth and development in power distribution systems. A large number of distributed generation units including renewable energy sources such as wind turbines, PV generators or fuel cells are being integrated into power systems at the distribution level. Distribution grids in open loop are typical, but now there is a rising trend towards closed loop topologies providing more power carrying capacities at distribution level. Although closed loop configurations improve power quality and the environmentally friendly nature of grid connected systems, they present some challenging issues at the same time. For instance, it is now possible to have power flows in many arbitrary directions instead of the traditional unidirectional power transfer. The penetration of DGs changes the traditional distribution power system short circuit power, fault current level, and the characteristics of the fault current, such as amplitude, direction, and distribution [1]. Protection of these distributed networks with the penetration of DGs is a focal point in the research of power systems and over the years there has been significant effort to improve its reliability by incorporating more sophisticated methods for fault detection, classification, and localization.

Power system faults are characterized as any abnormal change in system quantities like voltage levels, amounts of current, or frequency of the power signal. Mostly the occurrence of faults is a random event that may happen beyond the control of humans. Hence there is a need of an accurate, interconnected protection system that can detect the fault, identify its type, and then precisely locate the position of the fault in the power system. Sophisticated systems are already installed that take care of these faults and isolate the faulted zones from the rest of the power system.

The purpose behind power system fault analysis is to provide concrete information about the reasons that led to the interruption of supply, to restore the handover of power as soon as possible, and to perhaps minimize future occurrences if at all possible. Analysis should also provide sufficient understanding of the state of components of the protection system so that a set of preventive measures can be implemented to reduce the likelihood of service disruptions and equipment damage [2]. Fault detection, classification, and localization are a focal point in the research of power systems since the establishment of electricity transmission and distribution systems. Circuit breakers and other control elements are required to help protective relays to take appropriate action [3]. Fast detection of faults will have a significant impact on the equipment safety since it will engage the circuit breakers immediately before any significant damage occurs. Accuracy of fault location is not only significant for the clear reason of the timely repair and restoration of service, but it can also lead to identifying some specific location related faults. Hence, a longer term goal of preventing faults can be achieved. Over past several years, intensive research has been carried out in applied mathematics and digital signal processing to develop techniques for fault detection, classification, and localization in electrical power systems. Relaying and protection devices have been developed. Various signal processing schemes have been proposed that include principles of estimation, Artificial Neural Network [4], Wavelet transform [5] [6], Principal Component Analysis [7], Fuzzy logic [8], Support Vector Machines [9] and any combination of tools.

Most of the aforementioned works were validated via simulation studies, and their actual performance tested in a real-world scenario still remains unknown. This is because the hardware implementation of such fault protection techniques is rare.

In order to efficiently enhance the automation level in a power protection system, the main focus of this study is to develop a fast and accurate framework that can detect major types of faults at varying fault locations and fault resistances and classify them based on measurement samples of three-phase voltages and currents. One of the important aspects that this paper concentrates on is the analysis of the distribution grid line's phase voltages and currents during various fault conditions with distributed generation and coupling between feeders. This analysis is based on several assumptions concerning measurement nodes, fault locations, fault resistances, distributed generation infeed, and operational modes in radial configuration. Once the fault has been detected, the fault classification module is triggered that performs online fault classification based on offline trained models developed under a suitable computational scheme. In order to evaluate this proposed scheme, grid models are developed in MATLAB with integration of distributed generation. Faults are triggered and measurements are recorded for training purpose. These models are then transferred to real-time simulator hardware for real-time evaluation. Physical measurements are taken directly from the hardware and fed into a Programmable Logic Controller which performs online detection and classification of faults triggered under predetermined scenarios and time durations.

Section II presents the system architecture of the proposed framework for fault detection and support vector machine based classification scheme. Simulation and evaluation of the framework for a typical distributed grid model is described in Section III. Section III also presents results of software analysis by using MATLAB/Simulink toolbox. Section IV presents experimental validation of results obtained after successful implementation of grid model simulation on real-time hardware and PLC based implementation of the proposed protection scheme.

II. PROPOSED FRAMEWORK

In this section, the proposed framework for SVM based fault classification is introduced from a systematic view. The framework involves two primary modules: 1) fault detection and 2) fault classification.

Specifically the first step of the framework detects the presence of fault in the power system in real-time. This involves sampling of three phase voltages and currents from any one of the equipped measurement units in the power distribution network during an occurrence of a fault event.

Then the method of symmetrical components is used to convert the three phase input signals to three sets of independent components, which are positive, negative, and zero sequences. Negative sequence component is calculated as a reliable indicator of the fault condition. [10]. If no fault is detected, the remaining part of the modules is not activated. On the other hand, if the detection module captures the feature of the fault, it will activate the fault classification module. The overall architecture of the proposed protection scheme is illustrated in Figure 1.

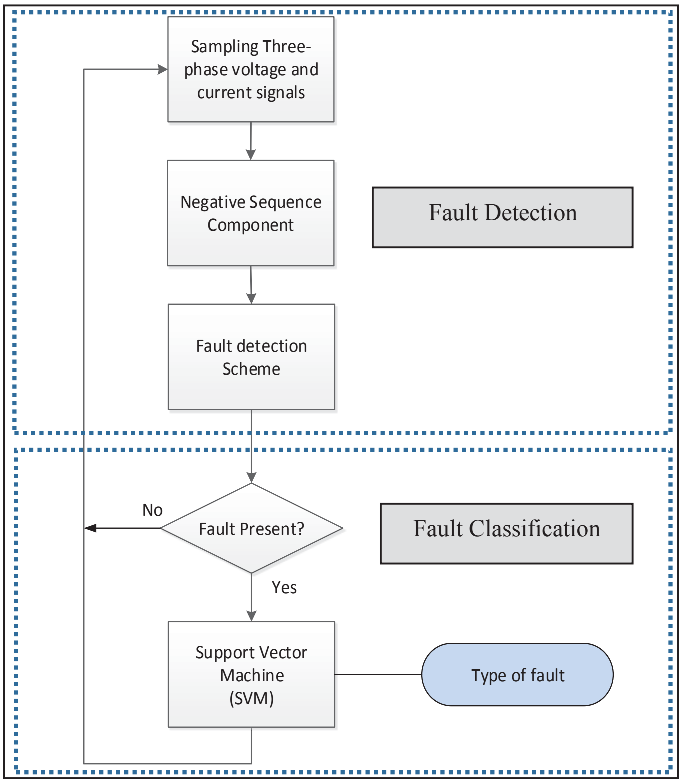

Figure 1: Proposed fault detection and classification scheme

A. Fault Detection Scheme

In electrical engineering, the method of symmetrical components is used to simplify analysis of unbalanced three-phase power systems under both normal and abnormal conditions. The analysis is simpler, because the resulting equations are mutually linearly independent if the circuit itself is balanced. [10]

In the proposed approach, the negative sequence components of voltages (v_{a2}, v_{b2}, v_{c2}) and currents (i_{a2}, i_{b2}, i_{c2}) are used as reliable fault indicators. NSCs of both voltage and current are calculated using the following equations

$$V_2(t) = \frac{1}{3}\left(V_a(t) + \alpha V_b(t) + \alpha^2 V_c(t)\right). \quad (1)$$

$$I_2(t) = \frac{1}{3}\left(I_a(t) + \alpha I_b(t) + \alpha^2 I_c(t)\right). \quad (2)$$

$$\alpha = 1\angle 120^{\circ} = e^{-j\pi\frac{2}{3}} = -0.5 + j0.866. \quad (3)$$

Figure 2 shows the schematic implementation of fault detection module

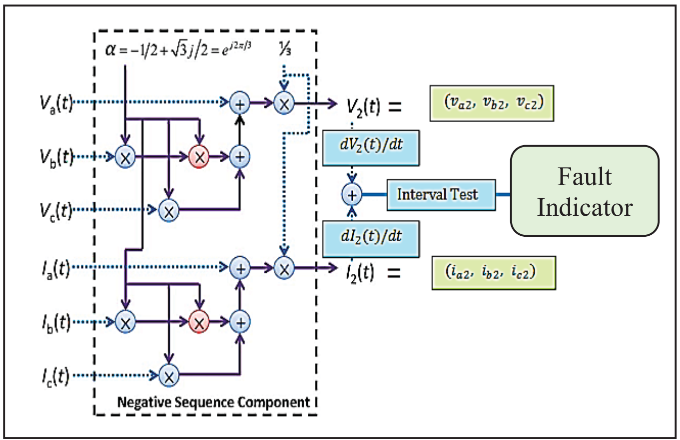

Figure 2: Schematic overview of fault detection module

However an unbalanced three phase power system suffers from large contents of negative sequence voltages and currents due to the unequal system impedances and unequal distribution of loads in each phase. In order to identify the variation in both non-zero negative sequence component V_2 and I_2, time differentiation is applied to both V_2 and I_2.

Finally, the differential negative sequence voltages and currents are summed together and the result undergoes an interval test with preset thresholds to perform a more stable detection of fault occurrence. It is worth mentioning that the fault detection scheme is also robust against frequency deviation and amplitude variation caused by the variation of operating parameters in the power system.

B. Multiclass Support Vector Machine using MATLAB

Support Vector Machine (SVM) is a novel intelligent machine self-learning method based on the statistical learning theory. The basic idea behind SVM revolves around mapping of n-dimensional inputs into a higher-dimensional feature space. In this feature space an optimal hyper-plane is determined that can classify input data points perfectly into two classes. The separating hyper-plane has the best generalization ability for the unseen data points. This optimal separating hyper-plane is generated by solving a constrained optimization problem.

A simple and effective example of linear classification with only two types of samples in a two dimensional space is shown in Figure 3. Suppose that we are given n training samples as $\{x_i, y_i\}$, $i = \{1\ldots n\}$, where $x_i \in R^d$ are input vectors and $y_i \in \{-1, +1\}$ which are class symbols with only two values. In the case of linearly separable data, the Hyper-Plane $f(x)$ makes a complete separation of two classes of samples, such that all samples with $y_i = +1$ fall on the right top side and where $f(x) > 0$. On the other hand samples with $y_i = -1$ fall on the left bottom side and there is $f(x) < 0$. All Hyper-Planes in R^d are parameterized by a vector w that is orthogonal to the Hyper-Plane, and a scalar b that is the bias of the Hyper-Plane with respect to origin.

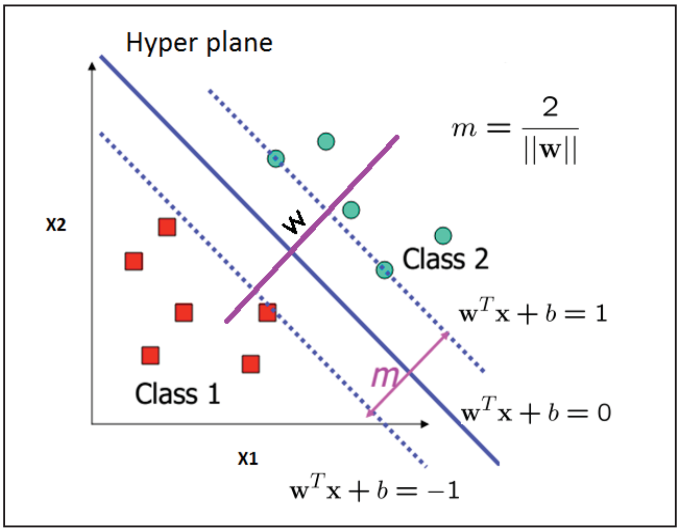

Figure 3: Optimal hyper plane

The optimal decision boundary can be found by solving the following constrained optimization problem also called primal formulation of SVM.

Minimize: $\frac{1}{||w||}$

Subject to: $y_i(w^T.x + b) \geq +1 \qquad \forall i$

The main method of SVM implementation is with Lagrange multipliers also called the dual formulation. This minimization problem is known as a Quadratic Programming Problem (QP) and it can be solved using many special purpose solvers [11]. One of the most commonly used solvers comes from Bioinformatics toolbox in MATLAB.

If the input data set is not linearly separable, there is a way to "pre-process" the data by defining a mapping $z = \phi(x)$ that transforms the d dimensional input vector x into a higher d' dimensional vector z. The mapping function $\phi(x)$ in higher dimensional feature space is called the *kernel trick.*

In order to classify all four basic types of faults, a multiclass support vector machine based on One-Versus-All approach [12] was developed. The multiclass SVM is comprised of four individual binary class SVM models for each type of fault. The detailed classification functions and the output patterns of these SVMs are summarized in TABLE 1.

TABLE 1: CLASSIFICATION FUNCTIONS AND THE OUTPUT OF SVM

Phase	Classification function	Detailed info	Output
Phase A	SVM for Phase to ground fault	A-G	1
	SVM for Double phase to ground fault	AB-G	2
	SVM for phase to phase fault	AB	3
	SVM for Three phase fault	ABC	4

Phase B	SVM for Phase to ground fault	B-G	1
	SVM for Double phase to ground fault	BC-G	2
	SVM for phase to phase fault	BC	3
	SVM for Three phase fault	ABC	4
Phase C	SVM for Phase to ground fault	C-G	1
	SVM for Double phase to ground fault	AC-G	2
	SVM for phase to phase fault	AC	3
	SVM for Three phase fault	ABC	4

C. Fault Classification Scheme

Once a fault has been detected by the fault detection module, the classification module is automatically triggered. The fault classification module takes a six dimensional input vector comprising three phase voltage and phase current samples. The vector can be expressed as

$$x = [V_a, V_b, V_c, I_a, I_b, I_c]. \qquad (4)$$

There are four major types of faults considered in this study

- Signal phase-to-ground faults : A-G, B-G, C-G;
- Double phase-to-ground faults : AB-G, BC-G, AC-G;
- Phase-to-phase faults : AB, BC, AC;
- Three phase fault : ABC;

Figure 3 shows the flowchart of SVM based fault classification scheme where each support vector machine designates a fault type. The individual SVM uses the input vector and calculates the functional value which results in either a positive or a negative output.

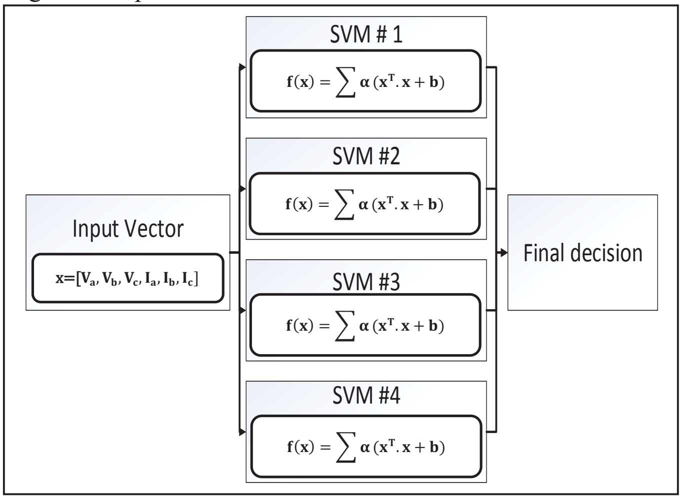

Figure 3: Flowchart of SVM based fault classification scheme

In order to classify all four major types of fault a multiclass decision based on *One-Versus-All* approach is used that provides the confidence of fault classification on each fault event.

III. Simulation and Evaluation of SVM based Fault Classification

This section presents the MATLAB based simulation of various fault types in a medium voltage grid model in integration with distributed generation. Different fault types at different fault locations are simulated and correspondingly their associated three-phase voltage and current measurements are recorded to build a training data base for the classifier. Once training is completed offline, testing is performed by simulating different types of faults and their classification is evaluated for true and false cases.

A. Medium Voltage Distribution Grid Model

Figure 4 shows the schematic diagram of a reference medium voltage distribution grid model which was tested under various star point topologies.

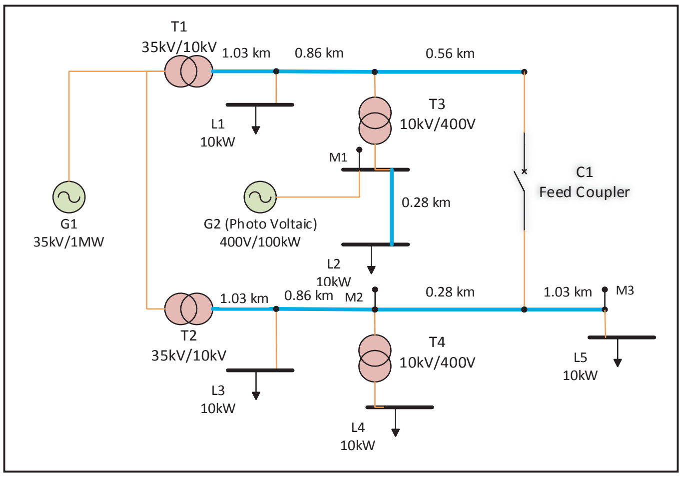

Figure 4: Schematic diagram of grid model with SS and DG

From Figure 4 G1 denotes a three-phase power source with rated voltage of 35 kilo-Volts and power capacity of 1 mega-Watts. The output from this source, used as infeed to the distribution grid is stepped down to 10kV on two separate substations denoted by T1 and T2 respectively. Feeder 1 contains a substation at 10kV denoted by T1 and a secondary substation at 400V at a distance of 1.89 km denoted by T3. A Photovoltaic based distributed generation G2 rated at 400V/100kW is integrated on this secondary substation. Moreover it also contains one fault inception and measurement block connected on secondary substation at M1. Feeder 2 contains a substation at 10kV denoted by T2 and a secondary substation at 400V at a distance of 1.89 km denoted by T4. It has backward as well as forward fault inception and measurement blocks to simulate both type of faults in the presence of closed loop operation of the grid. They are denoted by M2, and M3 respectively. There are in total three measurement points integrated in the grid on each fault inception block. Two of them in feeder 2, both on substation T1 and are denoted by M2, and M3; while the third measurement point is in feeder 1, integrated on secondary substation T3 denoted by M1. All measurements points provide per unit values of three-phase voltages and currents samples. A feed coupler is modelled as a switch that can be controlled with the prime purpose of simulating the ring topology of a distribution network. It is denoted by C1 in the schematic. When the feed coupler is activated, both distribution feeders are connected to each other.

B. Training and Testing Scheme for SVM

Figure 6 shows the flow chart for offline training and testing of support vector machine application as fault classifier.

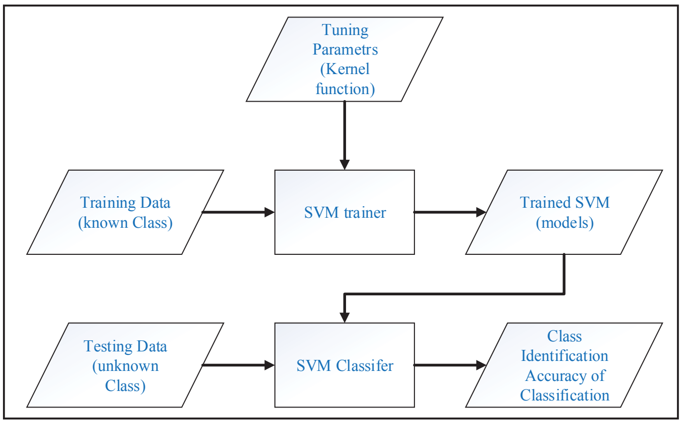

Figure 6: Flow chart for offline training and testing of SVMs

The training phase includes selection of parameters associated with the kernel function [11] and preparation of input data with known class types. Offline training is performed that generates models of SVM related to each class type. It should be noted that the size of training data, the dimension of inputs, and the selection of kernel function has a direct impact on the performance of the SVM classifier in terms of convergence and related time consumption for both training and testing phases.

The testing phase is also performed offline for simulation purpose. This requires input of data with unknown class to the SVM classifier. The classifier uses models that are generated from training phase to predict the class type associated with each testing case. As a result the accuracy of classification can be analyzed by generating cross validation data from training samples.

C. Results on Fault Classification

TABLE 2, TABLE 3 and TABLE 4 shows cross validation statistics of different fault cases against offline trained and tested samples based on RMS values of voltage and current signals. A Radial basis kernel function is used in each training phase with a different tuning parameter σ.

TABLE 2: FAULT CLASSIFICATION STATISTICS: σ=1

Fault Type	Test Cases	A-G	A-B	AB-G	ABC	Hit Rate
A-G	3000	2123	654	198	25	70%
A-B	3000	412	1567	911	110	52%
AB-G	3000	512	311	1821	356	60%
ABC	3000	138	223	625	2014	67%

TABLE 3: FAULT CLASSIFICATION STATISTICS: σ=0.75

Fault Type	Test Cases	A-G	A-B	AB-G	ABC	Hit Rate
A-G	3000	2670	110	178	42	89%
A-B	3000	333	2510	123	34	83%
AB-G	3000	210	153	2598	39	86%
ABC	3000	183	52	62	2703	90%

TABLE 4: FAULT CLASSIFICATION STATISTICS: σ=0.5

Fault Type	Test Cases	A-G	A-B	AB-G	ABC	Hit Rate
A-G	7000	7000	0	0	0	100%
A-B	7000	34	6954	12	0	99%
AB-G	7000	93	8	6899	0	98%
ABC	7000	0	40	0	6960	99%

IV. Real-time Validation of SVM Fault Classifier

In order to validate the simulation results in real-time, the grid model must be tested on hardware platform for different fault cases. For this purpose extensive hardware in the loop simulations are carried out using a real-time hardware simulator running the grid model under online induction of different fault conditions. Real-time fault detection and classification is implemented in hardware on a Bachmann M1 PLC system.

A. Real-Time Fault Detection and Classification Scheme

Figure 7 shows the scheme implemented in hardware for online fault detection and classification.

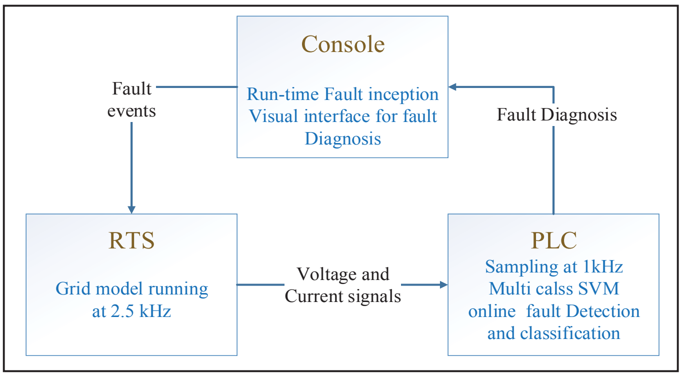

Figure 7: Hardware implementation scheme for online fault detection and classification

B. Opal-RT Real Time Simulator

An OPAL-RT real-time Hardware in the Loop (HIL) simulator at the disposal of ie3 lab at TU-Dortmund is used to run the grid model in real-time. The availability of multicores on this FPGA based real-time simulator is exploited in such a way that the complete grid model runs on one master core in conjunction with the PV based distributed generation on a slave core. This model splitting was necessary to avoid the introduction of overruns if the complete model was to run on a single core. The model runs on a step time of 400 micro seconds. A console block is used at the same time to generate different fault types on two forward fault locations and one backward fault location. It is important to note that at one time only one fault type is allowed to be triggered on an exactly one fault location. Meanwhile at any point in time the coupler may be activated to close the two grid feeders and enable the radial distribution network topology.

C. Hardware Setup

An overview of the hardware used at the disposal of ie3 lab is shown in Figure 8

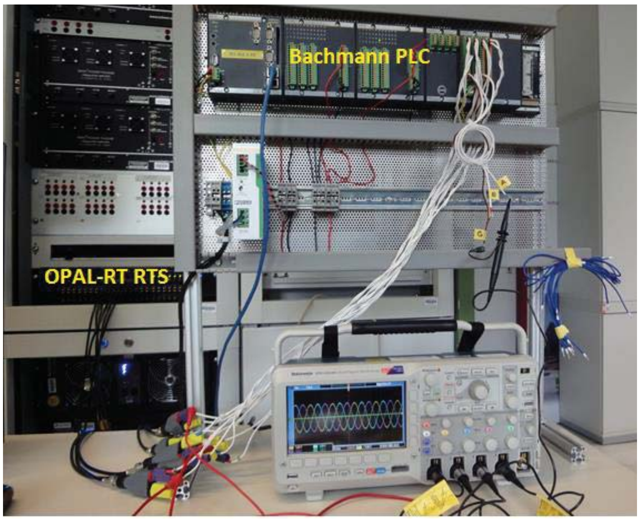

Figure 8: Experimental Setup

V. Discussion and Conclusion

A vital attribute of an electrical power network is the continuity of service with a high level of reliability. This has motivated many researchers to investigate power systems in an effort to improve power system reliability by focusing on fault detection, classification, and localization. We have presented a novel fault protection framework based on support vector machine that detects and classifies major fault types in electrical power systems. The proposed scheme has integrated fault detection and classification. The simulation results show that this algorithm has high accuracy in the classification performance and a wider generalization ability by using the learning and testing patterns in the voltage-current domain. A test distribution grid model with photovoltaic based distributed generation is successfully implemented in real-time on real-time simulator hardware. Experimental results show the hardware in loop simulation and the evaluation of the designed protection scheme implemented on a Programmable Logic Controller. The protective framework is of general applicability such that it can be deployed at any measurement point in the distribution grid. Hence it is shown that the algorithm has ability to correctly detect and classify fault occurrences making it suitable for modern distribution grids rendering high level of reliability.

While this study has a number of assumptions to secure the framework's successful implementation, there are several power network practices that can be included in possible future work to allow for improvements. These may include Integrating FACTS (Flexible Alternating Current Transmission System) in the distribution grid. The presented framework can be extended for fault localization. Instead of using offline trained models for fault classification, the framework can be adapted to perform online training making it more robust and immune to noise sources. The current scheme in combination with other fault classifiers can be hybridized to use the advantages of each protection scheme together. This would make it more adaptive to real world environments, especially when the work scope is expanded to include variation in system's parameters such as fault resistances, fault inception angles, high impedance faults, and ground level rise.

References

[1] a. W. K. Z. Chen, „Protection Coordination Based on Multi-agent for Distribution Power System with Distribution Generation Units," 2007.

[2] C. C. L. J. M. a. L. E. S. M. Kezunovic, „Automated fault analysis," IEEE Power Engineering Society, 2000.

[3] I. M.Kezunovic, „Detect and classify faults using neural nets," *Computer Applications in Power, IEEE,* Bd. 9, Nr. 4, Oct 1996.

[4] D. L. L. A.K. Ghosh, „The Classification Of Power System Disturbance Waveforms Using A Neural Network Approach," *IEEE Transactions on Power Delivery,* Bd. 10, Nr. 1, 1995.

[5] K. G. a. N. Kumarappan, „Comparative Study of Fault Identification and Classification on EHV Lines Using Discrete Wavelet Transform and Fourier Transform Based ANN," *International Journal of Electrical, Computer, and Systems Engineering,* 2008.

[6] S. G. Mallat, „A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transanctions on Pattern Analysis Machine Intelligence,* Bd. 11, Nr. 7, p. 674–693, 1989.

[7] Q. H. Alsafasfeh, „Fault Classification and Localization in Power Systems Using Fault Signatures and Principal Components Analysis," *Energy and Power Engineering,* Bd. 4, 2012.

[8] S. M. M. M. A. S. a. A. C. L. P. K. Dash, „Classification of Power System Disturbances Using a Fuzzy Expert System and a Fourier Linear Combiner," *IEEE transactions on power delivery,* Bd. 15, Nr. 2, 2000.

[9] P. S. S. &. P. G. Dash, „Fault classification and section identification of an advanced series-Compensated Transmission Line Using Support Vector Machine," *IEEE Transactions on power delivery,* Bd. 22, 2007.

[10] H. Saadat, Power System analysis, McGraw Hill, 2002.

[11] O. C. D. D. a. J. W. E. L. Bottou, Large Scale Kernel Machines, MIT: Cambridge, MA: MIT Press, 2007.

[12] B. T. D. &. K. H. Ravikumar, „Intelligent Approach for Fault Diagnosis in Power Transmission Systems Using Support Vector Machine," *International Journal of Emerging Electric Power Systems,* 2007.