Feature Extraction and Signal Classification of Traveling Wave Reflection Patterns for Line Fault Location

Haris Rehman and Björn Keune

***Abstract*—A vital aspect in the operation of electrical power systems is the fast and accurate locating of line faults for an effective isolation and clearance process. Reflection patterns of fault induced traveling waves provide valuable information about the fault location but are too complex for manual analysis. Instead methods from machine learning can be applied for pattern classification. However, it must be answered which features can be extracted from the reflection patterns and which of those features have the strongest impact on the overall classification performance. In this paper the method of Principal Component Analysis (PCA) is proposed in order to evaluate the impact of various features and their combinations on the pattern classification with Support Vector Machines (SVM) in advance. The proposed scheme is tested with simulated line fault data from the reduced IEEE 34 node test feeder. A comparison between feature selection and classification result confirms that PCA helps to discard irrelevant features and thus reduce the dimensionality of input data which in turn speeds up the classifier training process and takes less parameter storage capacity.**

***Index Terms*—Line Fault, Principal Component Analysis, Support Vector Machine, Traveling Wave**

I. Introduction

A vital aspect in the operation of electrical power systems is the fast and accurate locating of line faults for an effective isolation and clearance process [1]. In earlier times, this was achieved with the help of electro-mechanical relays and circuit breakers. Thereby, measured voltages and currents were converted into mechanical forces which operated the relays when a threshold limit was crossed [2]. Nowadays, the fault isolation principle is still the same but digital relays have widely replaced electro-mechanical solutions. However, in distribution level there are circuit breakers only available at the substations feeder bays and supply interruption is to be avoided in case of single line fault. Electrical faults are caused by human error, e.g. at construction sites, or by natural phenomena such as lightning strikes, trees falling into overhead lines or simply by age and its influence on isolation degradation or similar. The nature of fault depends on its cause and can be categorized, e.g. as transient or permanent and symmetrical or asymmetrical [3]. Studies have shown that in 80 % of the times these faults are line to ground faults. Different techniques stand available for fault locating based on the measurement of voltages and currents. Thereby, the accuracy of fault locating is one main aspect in on-going research. Using traveling waves (TW) propagation in protection was first proposed by Dommel and Michels in 1978 to detect and locate transmission line faults [4]. Although TW based schemes may provide both a high fault detection sensitivity and an accurate fault location there are some imperfections in them. For instance difficulties arise for the faults near the bus bars or for those occurring at near zero voltage crossing which leads to rather small wave amplitudes. Moreover, the wave propagation is affected by the system parameters and the network configuration. Eventhough, it is possible to determine the fault location with a single measurement node only. Thereby, the reflection patterns of fault induced TWs provide valuable information about the fault location. Yet, those patterns are rather complex for manual analysis. Instead methods from machine learning may be applied for pattern classification. To use the complete reflection pattern as input data for training classifiers is rather unfortunate since it goes along with higher computational effort. Instead single features of a pattern measurement shall be used as input data. However, it must be answered which meaningful features can be extracted from the reflection patterns and which of those features have the strongest impact on the overall classification performance.

In this paper the method of Principal Component Analysis (PCA) is proposed in order to evaluate the impact of various features and their combinations on pattern classification with the method of Support Vector Machines (SVM) in advance. The underlying objective is to investigate whether PCA can help to reduce the computational effort of TW reflection pattern classification and still provide accurate fault locating results. Thereby, a reduced IEEE 34 node test feeder is simulated to gain transient line fault data including the corresponding TW reflection patterns. Next, different features are extracted from those patterns and each ones significance is evaluated by PCA in order to find those which provide the most useful information about the fault location. Finally, those features proposed by PCA are cross-validated and the different partitions used as either training or test input for SVM. In the light of SVM results the benefits of PCA feature extraction for classification performance are evaluated. Thereby, the usefulness for discarding irrelevant features and thus reducing the dimensionality of input data are analysed. The complete work is based on simulations only since no TW measurements from a real power system are available to the authors that could be used for the investigation carried out.

Haris Rehman and Björn Keune are with the Technical University of Dortmund (E-Mail: haris.rehman@tu-dortmund.de, bjoern.keune@tu-dortmund.de).

II. Traveling waves

Faults occurring in the electrical power system generate TWs that propagate along the lines close to the speed of light [5]. Thereby, they travel back and forth as they get partly reflected and transmitted at points of discontinuity. The propagation paths can be illustrated by the Bewley-Lattice diagram [6] as shown in Figure 1 with two bus bars A and B and the corresponding TWs arrival times t_A and t_B.

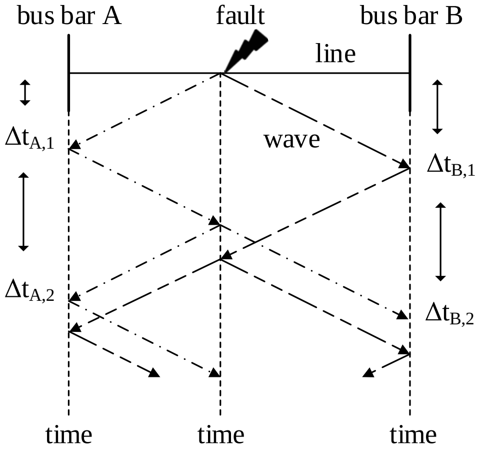

Figure 1: Common Bewley-Lattice diagram for line fault.

Due to the three phase power system the measured TWs have to be decomposed into their modal components first, namely α-, β- and γ-mode. Thus, with the help of Clarke's transform, which can be understood as symmetrical component calculation but for momentaneous values, each mode can be treated individually. In this work only the α-mode is used.

III. Proposed Scheme

The overall architecture of the proposed scheme for feature extraction and pattern classification is illustrated in Figure 2.

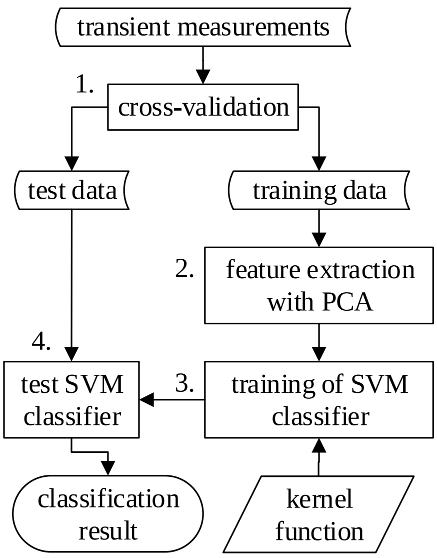

Figure 2: Feature extraction and fault classification scheme.

First, the transient measurements α-modes of simulated line faults are cross-validated to create training and test data partitions. Next, a pool of different features is derived from the transient measurements. From those features only the best are selected with the help of PCA. Following that, the selected features are forwarded as training input in different combinations for multiple SVMs. Thereby, the radial basis kernel function is used in each case. The trained SVM classifiers are then tested with the test data. Finally, the classification results are compared with the proposed feature selection of PCA.

IV. Distribution Grid Model

Figure 3 shows the reduced IEEE 34 node test feeder configuration used in this work. This reduction is due to the assumption of an ideal wave reflection coefficient of plus one for the transformer at node 814 which means that everything behind that node can be neglected. The voltage measurement M is placed at node 800. The nominal operating frequency is 50 Hz and the nominal distribution voltage level is 24.9 kV.

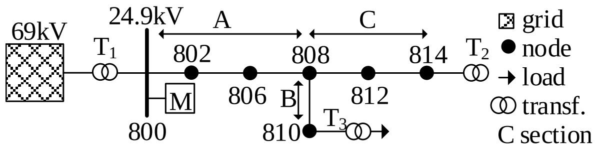

Figure 3: Reduced IEEE 34 node test feeder configuration.

Table 1 shows the electrical line parameters with the resistance r', the inductance l' and the capacitance c' as well as the corresponding wave velocities. The line lengths and grid section affiliations are given in Table 2.

Mode	r' (Ω/km)	l' (mH/km)	c' (nF/km)	v (km/μs)
α	0.136	0.908	12.43	0.298
γ	0.984	2.368	5.823	0.269

Table 1: Electrical line parameters.

Section	Node A	Node B	Line	Length (km)
A	800	802	L_1	0.786
A	802	806	L_2	0.527
A	806	808	L_3	9.824
B	808	810	L_4	1.769
C	810	812	L_5	11.43
C	812	814	L_6	9.062

Table 2: Line lengths.

V. Transient Fault Data Simulation

Different line fault scenarios are simulated in MATLAB/Simulink® for the reduced IEEE 34 node test feeder. Thereby, the model is simulated repeatedly with different fault locations, fault resistances and fault inception angles. The resulting three phase voltage measurements are recorded to build a training data pool for the SVM classifier based on the feature extraction with PCA.

A. Fault Scenarios

Table 3 lists the fault parameters used to create the different simulated scenarios for data generation. A single phase fault of random resistance value in the range of 1 Ω up to 1 kΩ and a random inception angle are simulated. For each simulation the fault location is moved 0.05 km along the feeder starting with line L_1 which means that twenty faults are simulated per km line length. The voltage measurement is placed at node 800.

Parameter	Description/Value
fault type	phase to ground fault
fault resistance	1Ω, 10Ω, 100Ω, 1000Ω
fault location	incremental distance of 0.05 km along complete feeder
fault time	three different timing of 0.0012 ms increment
measurement	three phase voltage at node 802

Table 3: Fault simulation parameters.

B. Signal Processing and Feature Extraction

The generated fault data is further used to extract meaningful features from the TW reflection patterns. Thereby, different methods of signal processing are used to prepare the data in time and frequency domain. Basically, with discrete fourier transform (DFT) and wavelet transform (WT) two different signal processing methods are used to prepare the transient voltage α-mode measurements. Afterwards, specific features are derived from the pre-processed reflection patterns.

- The DFT converts the N point input signal wave of time domain into a complex signal of frequency domain as defined (1). Thereby, n denotes the sample number and x its value. N is the total number of samples and k is the investigated frequency of the signal.

$$\underline{X}_k = \frac{1}{N}\sum_{n=0}^{N-1} x_n \cdot e^{-2\pi j \cdot \frac{nk}{N}} \tag{1}$$

- The WT is nowadays one of the most popular time-frequency transforms for transient signal analysis since it provides a very high time resolution in the high frequency range. In (2) its continuous form is given, where a is the scaling constant, b the translation and ψ is used the mother wavelet.

$$CWT(f,a,b) = \frac{1}{\sqrt{2}}\int_{-\infty}^{\infty} f(t)\psi\frac{(t-b)}{a}dx \tag{2}$$

Figure 4 shows a simulated voltage α-mode affected by single phase fault and Figure 5 its corresponding WT.

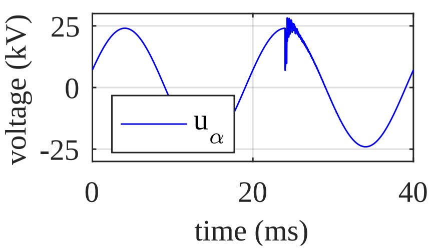

Figure 4: Transient fault record of voltage α-mode.

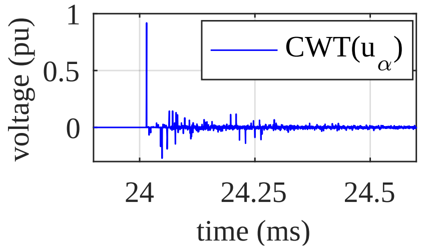

Figure 5: Wavelet transform of transient fault record.

In the following the different features that are extracted from the reflection patterns are described.

1) Power of Signal: The power content of electrical signals is considered as one most prominent feature and is calculated as in (3). Thereby, N denotes the number of samples and x(n) is the discrete samples of the signal.

$$p_x = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\|x(n)\|^2 \tag{3}$$

2) Number of Reflections: When a TW reaches a discontinuity such as a lines termination it gets reflected and travels back to where it came from. The electric fault is also a discontinuity and affects the number of arriving reflections at the measurement site within a finite time window. As the farthest distance in the grid model is 31 km approximately so the number of reflections S is calculated till the first reflection is expected to arrive from that farthest fault point as in (4). Thereby, $n \in [n_a, n_b]$ defines the considered time window with the sample of the first arriving TW n_a and the last considered reflection n_b.

$$S_{x_n>0} = \sum_{n=n_a}^{n=n_b} 1 \tag{4}$$

3) Average Timing of Arriving Reflections: Not only the number of arriving but also the average arrival time can be used as a feature indicating the fault location. Thereby, instead of each single TW arrival time the average of time differences T over the considered measurement window is calculated with (5) where t_n is the arrival time of a TW with the sample number n and S the total number of TWs.

$$T_{x_n>0} = \frac{1}{S_{x_n>0}}\sum_{n=n_a}^{n=n_b-1} t_{1+n} - t_n \quad \forall n \in x_n > 0 \tag{5}$$

4) Energy of Signal: Another feature to be taken into account is the fault signal energy. However, the signal energy changes with the TWs amplitudes that are also effected by the fault resistance. Instead of the energy only an energy equivalent E_x of the signal x may be calculated as in (6).

$$E_x = \sum_{n=0}^{N-1} x_n^2 \tag{6}$$

The signal energy can be either calculated from the original measured signal or from its WT.

5) Characteristic Frequencies: Another considered feature is the existence of path characteristic frequencies. To gain a simple corresponding scalar value the mean frequency of the signal over the frequency spectrum is calculated.

6) Wavelet Transform: The energy of the measured signal's WT is also considered as a feature similar to the approach of 4. Thereby, the voltage α-mode obtained from simulation is first decomposed by WT. Next, the energy equivalent is calculated from the wavelet detail's coefficients.

C. Feature Selection with Principal Component Analysis

In this research work the use of PCA is proposed in order to select the most suitable features as input for the SVM classifier. PCA is a multivariate technique that analyses the data table in which observations are described by several inter-correlated quantitative dependent variables [7]. Thereby, it represents its input data as a set of new orthogonal variables called principal components (PCs). Mathematically, it depends upon the eigenvalue decomposition of positive semi-definite matrices [8]. First step to the PCA is to obtain the covariance matrix for the given data by using the equation (7).

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{7}$$

Thereby, x is the input vector and n is the number of elements of the vector x. $\bar{X}$ indicates the mean of the set and with the mean of the data standard deviation has been calculated to find the covariance matrix. The covariance matrix is then calculated as in (8).

$$cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{(n-1)} \tag{8}$$

For example the covariance matrix C for an imaginary three dimensional data set with the input vectors x,y and z can be calculated with (9).

$$C = \begin{Bmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,y) & cov(z,y) & cov(z,z) \end{Bmatrix} \tag{9}$$

As the PCs are based upon the calculations of eigenvalues so from the above covariance matrix its eigenvalues can be calculated with (10). Thereby, C is the covariance matrix, I the identitiy matrix and λ the eigenvalues.

$$(C - \lambda I) = 0 \tag{10}$$

The highest value of λ is the first PC, its second highest value is the second PC and so forth and so on.
As described in chapter V-B six features are derived from the original fault data and analysed with the help of PCA. Figure 6 shows the results of PCA as a variance graph which indicate the impact of each single feature on pattern classification performance. Those PCs with a high variance are likely to make a great impact whereas the other ones are likely to have only a minor influence as they contain less useful information about the line fault location. PCA changes the order of those original six features and arranges them into the sequence of PCs from high to low as concerns their significance. Table 4 lists the relation between the six analysed features and their corresponding PCs or rather their likely priority order for pattern classification. Since one of the features has no considerable effect PC number six is not shown in the variance graph.

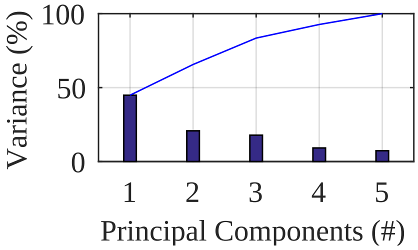

Figure 6: Principal components and their variances.

Feature (#)	Description	PC (#)
F_1	power of signal	PC_2
F_2	mean frequency of signal	PC_5
F_3	wavelet energy of signal	PC_3
F_4	average timing of reflections	PC_6
F_5	energy of signal	PC_1
F_6	number of arriving reflections	PC_4

Table 4: Features and their principal components.

The first three PCs have almost 65 % variance together and are likely to offer good classification results regarding the fault location. Their effect on the original data is illustrated in the 3-d-plot of Figure 7.

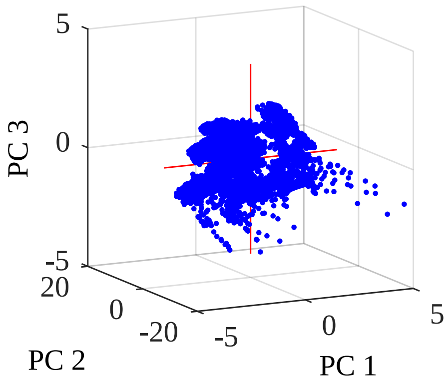

Figure 7: 3-d-plot of first three principal components.

D. Fault Classification with Support Vector Machine

After the impact of those considered six features is evaluated with PCA the most useful features of the training data are selected as input for fault classification with SVM in order to determine the fault location of the test data. The basic principle of SVM is to determine the optimal hyper plane h in the d-dimensional feature space so that each input data $x_i \in R^d$ can be assigned to its class y_i. The hyper plane h is commonly defined by ts normal w and bias b as in (11). To achieve this the dual Lagrange function $L(w,b,\alpha)$ of (12) shall be solved with the norm w, the bias b, the support vectors α, the classes y, the number of input data n and the kernel function K.

$$h := \{x \in R^d | \langle w, x \rangle + b = 0\} \tag{11}$$

$$L_D(w,b,\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \tag{12}$$

Thereby, the constraints of (13) are to be fulfilled.

$$\alpha_i \geq 0 \cap \sum_{i=1}^{n} \alpha_i y_i = 0 \quad \forall i \tag{13}$$

For the training part different feature combinations are used as input. For each combination one SVM model is trained with Radial Base Function (RBF) kernel. Thereby, it is taken into consideration that the training data, the dimension of inputs and the kernel function impact the overall performance of the trained model. For the sake of comparison each feature combination is treated with the same RBF kernel function. For the testing part the trained model is used to predict the class or rather the line fault location which can be either the faulted line or the fault distance from the measurement site. As mentioned before cross-validation with a ratio of 90 % training to 10 % test data was applied to create suitable partitions from the original data. Finally, the classification results are compared to the expectations of PCA feature selection.

1) Classification Reference Results: Table 5 shows the classification results of three trained SVM models with all six features as input. Thereby, the classification aims to determine the faulted feeder section. The results are to be reference for the classification results with only selected features of PCA.

SVM Model No.	1	2	3
Used Features	F_1-F_6	F_1-F_6	F_1-F_6
No. of Test Cases	10700	10700	10700
Correctly Classified	9941	9933	10263
Mis-Classified	759	767	437
Positive Results (%)	92.90	92.83	95.91

Table 5: Reference classification results.

2) Classification with Features Combination: Table 6 lists the considered feature combinations and average classification results. Thereby, both promising and less promising features are mixed. With the combination of feature one (F_1) and feature five (F_5) approximately the same classification results can be achieved as with all features combined. This confirms the expectations based on the feature selection with PCA as listed in Table 4. The naming sequence of the features in Table 4 and Table 6 is exactly the same which means F_1 is in both cases the signal power whereas F_5 is its energy. Indeed, the nature of these two features is very similar for dicrete-time signals.

Comb.	F_1	F_2	F_3	F_4	F_5	F_6	Avg. Result (%)
1	*				*		92.69
2			*		*		83.14
3	*		*				90.27
4		*		*			75.38
5				*		*	77.01
6			*			*	72.23
7			*	*			63.22
8					*	*	70.30
9		*	*				65.72

Table 6: Average combination classification results.

Since the TW propagating forth and back along the most direct path between fault location and measurement can be assumed to have the highest energy in the signal's frequency spectrum this might be one possible explanation why those two features deliver the best results. Because in the end, the fault location and resistance determine both the rate and ratio of reflected and transmitted TWs.

Figure 8 shows the SVM classification results for faulted section locating in detail where each group of bars represents a feature combination as listed in Table 6. For each combination the cross-validation, training and test is repeated ten times. As expected the classification results correlates with the quality of the selected features as input for the classifier.

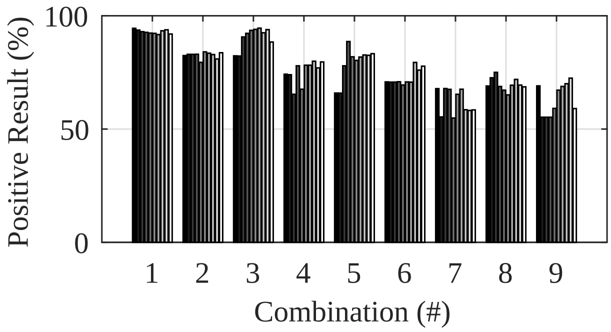

Figure 8: Results of two features combinations.

Furthermore, SVM models are trained for the combination of four features as listed in Table 7. Thereby, the training and test process is repeated four times as shown in Figure 9. The results confirm that eventhough a consideration of another two features does indeed improve the classification quality but that the combination of the two best features as selected by PCA is only barely surpassed. However, there are some downsides to the consideration of additional features since it increases the dimension of the SVM model which means longer training periods and more required storage capacity.

Comb.	F_1	F_2	F_3	F_4	F_5	F_6	Avg. Result (%)
1		*	*	*		*	89.78
2	*	*	*			*	93.15
3	*		*		*	*	79.43
4	*		*	*		*	95.76

Table 7: Four features combinations.

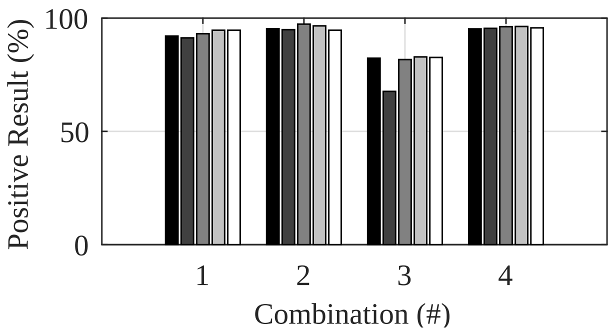

Figure 9: Four features combinations results.

Now that with the help of PCA the faulted grid section can be accurately identified the question arises if it is possible to both classify the fault type in case not only single phase faults are considered and locate the faulted line with the same good two features as input. Regarding the fault location classification, a distance offset of two hundred up to three hundred meters is still considered as an acceptable result. Indeed, a correct fault type classification of approximately 95 % and an acceptable fault distance of approximately 80 % are achieved with those same two features. Thereby, correct fault type classification means that among different fault types a single phase fault is correctly identified so that the trained SVM can be applied for the fault locating process.

It must be noted that only limited practical experiences regarding the accuracy of fault locating schemes with TWs exist within the research community which makes it difficult to estimate the achievable accuracy of the presented method. Up to now, such applications are only known from transmission level with line lengths of up to one hundred kilometres and two-terminal time-synchronized measurements. However, it can be safely stated that the accuracy with only a single measurement as presented in this work will be less accurate. Still, due to the shorter line lengths in medium voltage level the measured waveform will be less affected by frequency dependent attenuation and dispersion which make TW peak detection in transmission level such a technical challenge. Furthermore, the focus of this work is to investigate if PCA might be a suitable method in order to extract meaningful features from TW reflection patterns and thus save computational effort instead of using many features for classification with minor impact on the overall classification result.

VI. Conclusion

In this research work it is proposed to make use of PCA for the selection of suitable features from the local measurements of TW reflection patterns in order to use those features as input for a SVM classifier with the objective of fault locating. The classification test results confirm the proposal since only two selected features with the help of PCA determine the overall classification quality with a correct classification of more than 90 % whereas an additional consideration of the other features has only a minor impact. This proposed approach is considerably time saving and needs less computation effort compared to the other machine learning or complex deep learning techniques when combined with PCA. Thus, PCA significantly reduces the computational effort for classifier training. Indeed, it was possible to decrease the training time. The SVM classifiers were trained on a personal computer with 16 GB of RAM and intel core processor. For the good features selected by PCA the SVM classifier took 1.77 minutes to train whereas the total of six features as input took 19.52 minutes. After seeing this comparison between time and accuracy, it is safe to point out that PCA contributes to the application of machine learning in power system protection as it helps to extract significant features from the process values which in turn improves the fault classification process.

In future works not only star point grounding in isolated mode should be considered but also resonant grounding which is more common in medium voltage power systems, especially in central Europe. Furthermore, TW measurements should be taken from a real power system for final validation of the proposed scheme.

Literatur

[1] M.M. Polycarpou and A.J. Helmicki: *Automated fault detection and accommodation: A learning systems approach*. In: *IEEE Transactions on Systems, Man, and Cybernetics* 25 (1995), Nr. 11, S. 1447–1458

[2] Brown, Matthew H and Sedano, Richard P: *Electricity transmission: a primer*. (2004)

[3] Tleis, Nasser: *Power systems modelling and fault analysis: theory and practice*. Newnes, 2007

[4] T. Kawady and J. Stenzel: *Investigation of Practical Problems for Digital FaultLocation Algorithms Based on EMTP Simulation*. In: *In Transmission and Distribution Conference and Exhibition* 1 (2002), S. 118–123

[5] Wedepohl, LM: *Application of matrix methods to the solution of travelling-wave phenomena in polyphase systems*. In: *Proceedings of the Institution of Electrical Engineers* Bd. 110 IET, 1963, S. 2200–2212

[6] Datta, Brajagopal and Chatterjee, Saibal: *A literature review on use of Bewley's lattice diagram*. In: *Power and Energy in NERIST (ICPEN), 2012 1st International Conference on* IEEE, 2012, S. 1–4

[7] Wold, Svante and Esbensen, Kim and Geladi, Paul: *Principal component analysis*. In: *Chemometrics and intelligent laboratory systems* 2 (1987), Nr. 1-3, S. 37–52

[8] Abdi, Hervé and Williams, Lynne J: *Principal component analysis*. In: *Wiley interdisciplinary reviews: computational statistics* 2 (2010), Nr. 4, S. 433–459