On-site Multi-Class Feature Selection for Online Classification of Switchgear Actuations in the Distribution Grid

Christina Nicolaou[1,2], Ahmad Mansour[1] and Kristof Van Laerhoven[2]
[1] Robert-Bosch-GmbH, Robert-Bosch-Campus 1, 71272 Renningen
[2] University of Siegen, Adolf-Reichwein-Straße 2, 57076 Siegen

Abstract

The electrical grid is highly dependent on switchgear to maintain a safe and reliable power transmission. For this reason, the interest in on-site, non-invasive monitoring solutions including the detection of switch operations, their differentiation and ageing has significantly increased in the last years. Thereby, the research field of tracking acoustic emissions generated during the switching using low-cost micro-electro-mechanical system (MEMS) based sensors is emerging. This paper presents a computationally efficient method for selecting process- and design-specific features on-site (on a sensor system or gateway) to eliminate the need of prior offline training. This ensures generalized usability for different switch types and sensor positions without high re-training effort. The selected features are further used for online multi-class classification of switching processes. The proposed self-learning method, as well as the use of the MEMS sensors (acoustic and vibration), are both evaluated for classification performance on switchgear measurements during twelve different processes, leading to a robust classification with an accuracy of over 95 % in average.

1 Introduction

Switchgear plays an indispensable role for the reliability and resilience of power systems by protecting and de-energizing critical electrical equipment in the event of a fault. As the demands on the power grid become more complex (e.g., higher safety and environmental standards, reduction of power outages and improved power quality [1]), more grid transparency is needed at all voltage levels. While continuous online monitoring of switching operations including contact travel curve and timing, number of operations, electrical parameters as well as insulation condition is implemented as a standard in the higher voltage levels of the electrical network [1], this is only the case in sporadic exceptions in the medium-voltage (MV) distribution grid. Economically, large-scale condition assessment using cost-intensive monitoring systems, such as the ones operating in the high-voltage level, is not justified as the cost of the monitoring system are of a similar order of magnitude as those of the equipment under investigation. Nevertheless, knowledge of the number and frequency of switching processes is of fundamental interest as it decisively determines wear and aging effects of the MV switches, which are the cause of 42.3 % of major failures [2]. For this reason, cost-effective solutions with low installation effort and easy handling for the detection of switching operations are required. Great potential is promised by using low-cost micro-electro-mechanical system (MEMS) based acoustic and vibration sensors in combination with suitable intelligent data evaluation.

In this paper, we consider the MEMS-based monitoring of metal encapsulated MV switchgear units, which are manufactured according to a modular principle from individual cubicles. The cubicles consist of a variety of switches (e.g. load break switch, circuit breaker or earthing switch), fuses and auxiliary equipment (see **Figure 1**). While load break switches are used for (dis)connecting single assets or stations from the grid at rated current, earthing switches are used for grounding and can also be used in case of short-circuits. All switches are equipped with spring-assisted mechanisms to allow safe operations independent of the operational speed. As one switchgear unit usually contains more than one switch to be monitored, the unambiguous assignment of detected switching operations to the corresponding switch as well as the differentiation between switching processes (on and off) is important to be able to track the frequency and number of operations as well as aging effects of the individual breakers. Existing solutions rely on limit switches being triggered at the end of a switching operation or on the monitoring of electrical parameters to detect which circuit is opened or reconnected. As a retrofit solution, these approaches have the disadvantage of a high installation and wiring effort, since they need an installation inside the unit at each breaker. In contrast, our solution provides a way to detect the different switch operations non-invasively using only one MEMS-based sensor system. This greatly simplifies the installation and promises to provide a suitable solution for network operators as well as a widespread use.

Multi-class classification must be carried out to assign the recorded signals to the correct switch. To train their models, standard methods collect training data to extract features and select the application-specific ones that can separate classes in the feature space. After the offline training, the trained model is implemented for the use in the field e.g. on a gateway or a microcontroller. The repetitions of those offline learning cycles to learn design-specific features for the high diversity of switchgear types and manufacturers existing in the field lead to a considerable effort. Furthermore, a change in the mechanics due to component exchange or aging also requires repeating the procedure again. Those methods are therefore less suitable for assessing the condition of switchgear. For this reason, it is particularly important to use generalizable models for universal applicability.

In prior work [3], we presented an online configuration concept for automatic differentiation between switching-on and -off based on an adapted Silhouette Score. In this paper, we extend the approach to select relevant features to optimally distinguish between operations of all switches in a unit. Thereby, our approach allows training and appropriate feature selection to be performed automatically without

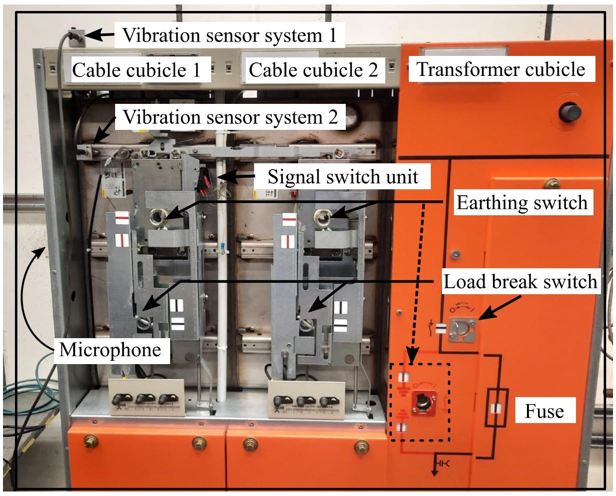

Figure 1 Part of a triple cubicle metal encapsulated SF6-gas-isolated MV switchgear unit from Driescher Wegberg with removed cable cubicle covers.

inference from the user, using only few labeled training data. The method is computationally efficient, so that the training can also be performed on-site (e.g., on a gateway or on the sensor system itself). Furthermore, the proposed method is not limited to switchgear classification but can be customized for other applications in the machine condition monitoring domain.
In the following, an overview of existing switchgear monitoring solutions is given before further introducing our proposed method for on-site multi-class classification. The approach is evaluated on switchgear measurements of one unit in a proof-of-concept study.

2 Related Work

The general trend from time-based to condition-based maintenance among grid operators leads to more and more attention being paid to the monitoring of switchgear. A comprehensive overview of condition assessment methods, their maturity and current research areas are given in [1, 4]. A promising mainstream research area is the vibration- and acoustic-based monitoring of switchgear, that has been studied since the 90s, mostly focusing on high-voltage circuit breakers [4]. Using acoustic signals generated during the switching processes promises to be a good solution for detecting mechanical failures, which are the origin of 52.6 % of major faults [2]. The key challenge for the algorithm development is the complexity of the acoustic signals, which results from very short switching durations with extremely wide energy bands in spectrum, strong non-linearity and non-stationarity, corruption by noise, dependency on the high variety of components and switching being only performed few times a year [4, 5, 6]. To distinguish between natural variations and (un)known fault conditions, appropriate features and classifiers have to be found [4, 7]. Main feature extraction methods are, on the one hand, classical spectral analysis based on the Fourier transform [8] and, on the other hand, approaches that are more suitable for non-stationary signals and extract local information in the time-frequency domain [5, 9]. To get frequency components in a good resolution without mixing effects, adapted multi-scale decomposition methods like empirical mode/wavelet decomposition, wavelet packet transform or variational mode decomposition are used [4, 5, 7]. Disadvantages of these solutions are, amongst others, the determination of the decomposition level, which might affect the analysis robustness and the considerable time to implement [7]. Furthermore, different classifiers to distinguish between normal and faulty equipment were reported in literature. An overview of methods with their (dis)advantages is given in [7, 9]. Thereby, these methods are mainly trained to detect certain mechanical failures (e.g., poor lubrication, spring fatigue or contact damage). However, they lack robustness, they highly depend on the training set as well as sensor position, and are limited in their generalizability given the diversity of switchgear types [4]. The high manual effort and expert knowledge to adapt these methods as well as the necessary precise sensors with high bandwidth are not justified at MV level, which is the reason why vibration monitoring and its usability for MV switchgear is still an open research topic [6]. A cost-effective, sensor-position-independent, and more generalizable method using self-learning algorithms and MEMS-systems is examined in the following. The method is primarily used to detect switch operations and assign them to the correct switch, though it can also be used to track trends in the extracted features due to ageing effects [3].

3 Multi-Class Feature Selection

In [3], we showed how ranked feature selection can be performed based on the largest information gain in a one-dimensional feature space using an adapted Silhouette Score for learning a binary classifier. Preconditions for the online classification are the known number of classes and few training data, which are collected on-site. With the adapted Silhouette Score, a feature quality score (f_C) can be calculated using the distance between features extracted from the training data and their own cluster center c_i and the nearest cluster center c_k as follows:

$$f_C = \frac{1}{N}\sum_{i=1}^{N} \frac{d_{ik}-d_{ii}}{\max(d_{ik},d_{ii})},$$

whereby N is the number of classes, d_{ii} is the average distance of feature values of one cluster to their own cluster center c_i and d_{ik} is the average distance of those features to the nearest cluster center c_k (see **Figure 2**). Feature scores of nearly one lead to a good classification and can be used for the inference phase, where the corresponding

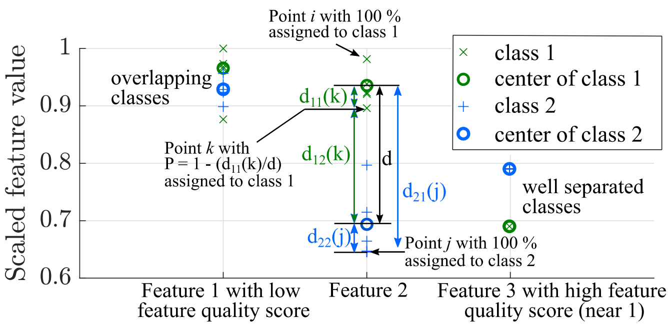

Figure 2 Example of the Silhouette Score to select suitable features as binary classifiers.

feature is extracted and its distance to the centers (converted into a class membership probability) is used as a classifier (see **Figure 2**). The extension of this method to multiple classes in the one-dimensional feature space can lead to strong overlapping clusters. To circumvent this, a multi-dimensional feature space is often created, which strongly distributes the features in space and thus leads to a better separability. However, with the increase of dimensions, the number of needed training data also increases exponentially (curse of dimensionality). This is critical as only few labeled training data can be provided for practical reasons. For this reason and as lower dimensions also lead to an easier interpretability, lower computational costs, and an increased robustness against overfitting, a transformation of the multi-class problem to a binary one is carried out. To get binary classifiers, two popular methods exist – one vs. all (OVA) and one vs. one (OVO). For OVA, m classifiers for m classes are learned, whereby each classifier distinguishes between the data of one class against all other $m-1$ classes. In contrast, OVO creates classifiers for all possible pairs of classes and therefore requires $m(m-1)/2$ classifiers, but yields better results [10]. For this reason, we propose a combination of the one vs. one classification with the adapted Silhouette Score. In our case, each operation (on/off) of a switch (earthing/load break switch) in the unit forms its own class. An example of each operation is shown in **Figure 3**, which illustrates that the characteristics of the signals differ depending on the cubicle, switch type and process. For the training, features are extracted from the training data of all switch operations. Features f_p from one class p are pairwise compared with the features $(f_l, f_m, f_n, ...)$ of the remaining classes $l, m, n, \ldots$ (one vs. one) by calculating the adapted Silhouette Score. The features that are best suited to separate these classes pairwise have a high Silhouette Score and are selected to serve as binary classifiers. In the inference phase, the selected features are calculated for every detected switching actuation and a selection between two classes is made for each classifier (**Figure 4**). Thus, given j classes, a decision can be made for one class up to $(j-1)$ times. Thereafter, the class with the most votes is assigned to the operation.

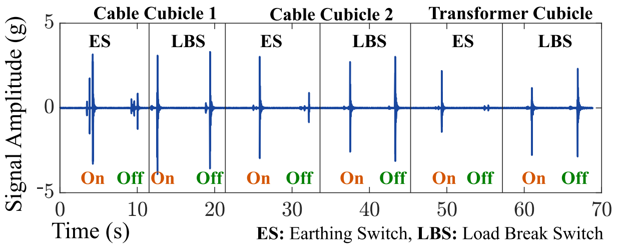

Figure 4 Example of switching processes for each switch

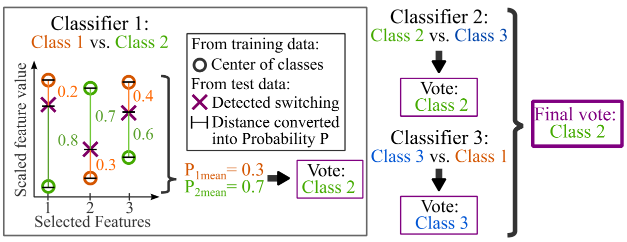

Figure 3 Combination of Silhouette Score and OVO for three classes and three features per classifier

4 Proof-of-Concept Study

In the following the benefits of our proposed method are evaluated on measurements of six switches gained on a MV switchgear unit in a proof-of-concept study.

4.1 Data

The data of this study were obtained at the MV switchgear unit shown in **Figure 1** using one MEMS-based stereo microphone (MBSM) with a sampling frequency of 48 kHz and two MEMS vibration sensor systems (VSS). The VSS consists of a single-axis high-bandwidth accelerometer (A1) with a sampling rate of 62.5 kHz and a second three-axis accelerometer (A2) with good resolution in low frequencies with a sampling rate of 2 kHz. The sensor systems were installed at fixed positions to examine their capability to distinguish between all actions of one unit. VSS1 was screwed near the earthing switch drive in cable cubicle 1 and VSS2 on top of the same cubicle. The MBSM was installed via a magnet approximately one meter behind the switchgear unit. Measurements of 200 complete on/off switching cycles were recorded at all six switches (load break switch and earthing switch in each cubicle).

4.2 Feature Selection and Classification

For the switch actuation detection and labeling of the training data, a threshold-based recognition is implemented as in [3], which leads to a detection accuracy of over 99 % and robustness against usual background noises (e.g., people speaking nearby). For the training phase, five recurrences of each process have been recorded and directly labeled with the process, the switch type, and the cubicle after sensor installation. For each classifier, the corresponding sensor signals are low pass filtered with different cut-off frequencies, as the switch system is characterized by a damping system with oscillations mainly lying in the low-frequency range. Afterwards, 21 different features (mean, variance, kurtosis, power, flatness, root mean square, absolute mean, maximum, minimum, dynamic range, crest factor, spectral centroid, median and dominant frequency) from the time and frequency domains are extracted and normalized. For each classifier, the best five features are selected with the adapted Silhouette Score to increase the robustness as in [3]. In the inference phase, the probability for every selected feature is calculated depending on the distance of the feature to the corresponding cluster centers learned during training (**Figure 2**). An average probability is built from the five selected features for each classifier and the corresponding class with the highest probability is selected. Afterwards, the class that was chosen most often by the classifiers is finally selected (voting). As the switches in one cubicle are safeguarded against each other, each cubicle can only be in one of three states – both switches are off, earthing switch is on and load break switch is off or the other way around. This knowledge can be incorporated to check the assignment or to decide which class is more plausible, if more than one class wins the same number of votes.

4.3 Results and Discussion

For the evaluation, five training measurements per switch and operation were randomly selected from all measured data and the classifiers are tested on the remaining 2340 operations. This process was repeated 10 times and the average accuracy per sensor was calculated:

Sensor	VSS1 A1	VSS1 A2	VSS2 A1	VSS2 A2	MBSM
Accuracy	96.2 %	93.5 %	91.8 %	95.5 %	97.5 %

For the differentiation between all processes and switches, all sensors lead to an accuracy between 92 % and 98 % with a standard deviation between 0.9 % and 1.9 % for the different training repetitions. The axis-orientations of the sensors play here a minor role. However, the sensor installed inside the first cable cubicle as well as the microphone tend to clip for measurements on the nearest switches, due to the large forces released when switching, thus exceeding the defined measurement range. For the microphone, a distance of at least 1-2 meters should be kept. Furthermore, both A1 sensors have the lowest SNR for the switching-off process of the earthing switches, which is better detectable for the accelerometers A2 with a greater resolution in the low frequency range. This is the case, as those switches do not release as much energy as the other switch processes. Furthermore, it should be taken into account that although the microphone can be used as a central sensor in the station for other use cases, such as intrusion detection, it is also susceptible to background noise, which could influence the switching detection. For the classifiers, the most frequently selected features are the dominant occurring frequency, mostly located below 1 kHz (supporting the approach of filtering the signal before feature extraction) and the absolute mean in the time and frequency domain, which is determined by the distance of the sensors to the individual switches. Furthermore, features learned for one sensor position cannot easily be transferred to another one, as the signal characteristics differ significantly. It is therefore important to be able to learn design-specific features locally, which supports our approach giving an opportunity for on-site self-learning.

4.4 Conclusions and further work

In this paper, we present an approach for on-site feature selection using an adaption of the Silhouette Score, which prevents time-consuming and application-specific offline training. These features are used for learning binary classifiers for an online one vs. one multi-class classification. We show that with this approach, twelve different switchgear actuations (that are automatically detected via a threshold) can be classified reliably. The suitability of two different MEMS-based sensor systems (acoustic and vibration) as well as two different installation positions are evaluated, and recommendations are provided. Further investigations include long-term testing of the method, where changing properties of the switch spring mechanics may lead to changes in the data distribution that need to be tracked (as in [3]) to potentially allow for re-learning. Additionally, the performance when the switches are not actuated for several weeks instead of several days as done here needs to be validated, and a state machine logic to further improve the classification accuracy can be implemented.

5 Acknowledgement

The project MAKSIM, on which this report is based, is funded by the Federal Ministry for Economic Affairs and Energy under the funding code 0350035B. The authors are responsible for the content of this publication. The authors would like to thank the Fritz Driescher KG for the cooperation and for providing access to their switchgears. Thanks goes especially to Mr. Bernards, Mr. Goertz and Mr. Kohnen for their help and support in the study and tests.

6 Literature

[1] N. Uzelac et al., "Non-intrusive methods for condition assessment of distribution and transmission switchgear," 2018.

[2] M. Runde et al., Final report of the 2004-2007 international enquiry on reliability of high voltage equipment, Paris: CIGRE Technical Brochure 509, 2012.

[3] C. Nicolaou, A. Mansour and K. Van Laerhoven, "On-site Online Feature Selection for Classification of Switchgear Actuations," - *arXiv:2105.13639,* 5 2021.

[4] A. A. Razi-Kazemi and K. Niayesh, "Condition Monitoring of High Voltage Circuit Breakers: Past to Future," *IEEE Transactions on Power Delivery,* vol. 36, p. 740–750, 4 2021.

[5] Q. Yang et al., "A New Vibration Analysis Approach for Detecting Mechanical Anomalies on Power Circuit Breakers," *IEEE Access,* vol. 7, p. 14070–14080, 2019.

[6] M.W. Hoffmann et al., "Integration of Novel Sensors and Machine Learning for Predictive Maintenance in Medium Voltage Switchgear to Enable the Energy and Mobility Revolutions," *Sensors,* vol. 20, p. 2099, 4 2020.

[7] S. Wan and L. Chen, "Fault Diagnosis of High-Voltage Circuit Breakers Using Mechanism Action Time and Hybrid Classifier," *IEEE Access,* vol. 7, p. 85146–85157, 2019.

[8] W. Wang et al., "Online monitoring of high-voltage switchgear installation," *The Journal of Engineering,* vol. 2019, p. 1238–1240, 3 2019.

[9] S. Ma et al., "High-Voltage Circuit Breaker Fault Diagnosis Using a Hybrid Feature Transformation Approach Based on Random Forest and Stacked Autoencoder," *IEEE Transactions on Industrial Electronics,* vol. 66, p. 9777–9788, 12 2019.

[10] P. C. Chaitra and R. S. Kumar, "A review of multi-class classification algorithms," *Int. J. Pure Appl. Math,* vol. 118, p. 17–26, 2018.