A SVM-based Setting of Protection Relays in Distribution Systems

Xiangtian Zheng, Xinbo Geng, *Student Member, IEEE,* Le Xie, *Senior Member, IEEE,* Dongliang Duan, *Member, IEEE,* Liuqing Yang, *Fellow, IEEE* and Shuguang Cui, *Fellow, IEEE*

***Abstract*—With deeper penetration of distributed energy resources (DERs) in the distribution system, the power flow patterns are becoming more diverse, rendering over-current relays in distribution systems with potential malfunctioning. This paper proposes a data-driven approach to design operating strategies of relays in distribution systems with high penetration of distributed energy resources (DERs). A support vector machine-based classifier for the relay setting is introduced. Compared with conventional over-current relays, the proposed method could operate more accurately with smaller false alarm rate. It is also robust to altering system conditions such as load levels and fluctuating DERs. Numerical studies based on the IEEE 123-node test feeder illustrate the potential of this approach.**

I. Introduction

A. Background

One of the most fundamental tasks in power systems operation is to keep the system secure during potential faults. With deeper penetration of renewables into both transmission and distribution systems, this task becomes increasingly difficult. Particularly, the expanding distributed energy resources (DER-s) reform the distribution grids and pose new challenges on the secure and reliable operations of distribution systems with high variability and limited predictability from renewables.

Relays are major devices protecting power system from faults. Typically, protection relays isolate faults from power grid and thus reduce negative impacts on other operating devices. Among a large variety of protection relays installed in distribution systems, over-current relays are the most widely-used ones. The operating strategy for conventional protection relays are often model-based. More specifically, conventional over-current relays (in short, *over-current relays*) operate according to a current-based characteristic curve, which determines whether to act and the waiting time if action is necessary. Typical over-current relays operate based on the magnitude of current and only have one threshold. Such operating strategy relies on two underlying assumptions: (1) the normal-state current is always smaller than the short-circuit current and (2) faults located closer to the relays cause higher measured short-circuit current. These two assumptions could be problematic with increasing penetration of DERs, fluctuating load levels and more complicated system topology. It is necessary to design better operating strategies for protection relays in the increasingly complex distribution system.

X. Zheng, X. Geng and L. Xie are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, 77840. Emails: {zxt0515, xbgeng, le.xie}@tamu.edu.

D. Duan is with the Department of Electrical and Computer Engineering, University of Wyoming, Laramie, WY. Email: dduan@uwyo.edu.

L. Yang is with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO. Email: lqyang@engr.colostate.edu

S. Cui is with the Shenzhen Research Institute of Bid Data, Chinese University of Hong Kong, Shenzhen, Guangdong, China. Email: shuguangcui@cuhk.edu.cn.

D. Duan, L. Yang and X. Zheng are visiting scientists and student at the Shenzhen Research Institute of Bid Data, Chinese University of Hong Kong.

B. Literature Review

There has been a substantial body of literature exploring advanced operating strategies for protection relays with an increasing amount of DERs into modern power systems [1]–[17]. Many of them (e.g. [3], [4]) incorporate machine learning techniques to improve the performance of over-current relays. Most studies of improving the over-current relays focus on the aspects of coordination [6], [7], fault detection [18] and fault section estimation [4], [18]. Among many machine learning algorithms, Support Vector Machine has shown its effectiveness and efficiency in many power system applications, e.g. electric market data analysis [19], [20] and wind forecasting [21], [22]. There are also a few references applying the SVM algorithm to design protection relays [9]–[13]. A closely related work is [9], where the authors propose intelligent relays based on SVM and hypothesis testing. Studies in [9]–[13], [17] demonstrate the effectiveness of SVM on protection relay designing, while most of them focus on the transmission grid. As suggested in Section I-A, expanding DERs leads to rising interest in designing intelligent relays in the distribution system. This paper is our first attempt to solve this problem.

C. Contribution of This Paper

Motivated by the need to make better use of the streaming data collected from increasing deployment of sensors and measurement devices in the distribution grid [23], this paper proposes a data-driven operating strategy for protection relays. Conventional over-current relays only utilize the information of current and require the prior assumption of the network topology, load level and generation level. The limitation of the information resources and the prior assumption determining the performance of over-current relays are usually imperfect in the increasingly complex distribution system. The proposed strategy has the following two major advantages over traditional relays: (1) it is purely data-driven and does not require any assumptions on the system topology or load levels; and (2) it makes better use of data. Two types of data in this paper (i.e. active and reactive power measurements) are considered,

978-1-5386-1006-0/18/$31.00 ©2018 IEEE

Authorized licensed use limited to: Texas A M University. Downloaded on December 17,2021 at 04:45:21 UTC from IEEE Xplore. Restrictions apply.

while the conventional over-current relays only utilize current measurements. Because of these two advantages, the proposed SVM-based relays could effectively reduce potential maloperation and misstrip rates.

The rest of this paper is organized as follows. Section II introduces the Support Vector Machine algorithm. Complete details on the proposed data-driven SVM-based strategy are provided in Section III. Section IV presents case studies and a comprehensive comparison between SVM-based relays and over-current relays. Concluding remarks and future works are presented in Section VI.

II. Support Vector Machine

This section introduces the Support Vector Machine algorithm, which lays the foundation of the proposed SVM-based relays in Section III.

A. *Classification Problem*

A *classifier* is an algorithm to determine the label y given a feature vector x. The feature vectors with the same label belong to the same *class*. A classification problem typically aims at finding the most accurate classifier. There are in general two categories of classifiers: parametric and non-parametric ones. The parameters of a parametric classifier are often obtained by fitting over a training set, which consists of both feature and label vectors.

Support Vector Machine is one of the most influential parametric classifiers. It was first proposed by Vapnik and Chvervonenkis in 1963, and the standard "soft-margin" formulation was proposed in 1992 [24]. It is built upon a very simple yet powerful idea: finding an optimal separating hyperplane with the widest margin as the optimal classifier. The remainder of this section introduces the SVM classifier in detail.

Note that we restrict our attention to the binary case in this paper. Namely, the label vector y can only take two values -1 and 1 ($y \in \{-1, 1\}$).

B. *Linearly Separable Case*

In this subsection, we assume the existence of at least one hyperplane separating two classes in the training dataset. The objective of SVM is to find the optimal separating hyperplane with the largest margin. Fig. 1 illustrates the basic idea of SVM.

Let $f(x) = w^\intercal x + b = 0$ depict a separating plane, then for the feature vectors x with label $y = 1$:

$$w^\intercal x + b \geq 1 \tag{1}$$

Similarly, for the feature vectors x with label $y = -1$:

$$w^\intercal x + b \leq -1 \tag{2}$$

Using the fact that $y \in \{-1, 1\}$, Eqn. (1)-(2) are equivalent to

$$y(w^\intercal x + b) \geq 1 \tag{3}$$

It is straightforward to obtain the width of the separating hyperplane (i.e. distance between $w^\intercal x + b = 1$ and $w^\intercal x + b = -1$) as $\frac{2}{||w||}$, then the SVM classifier is to solve the following optimization problem:

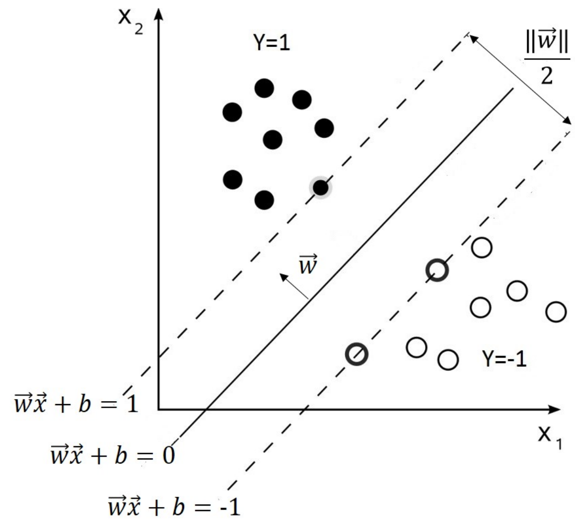

Fig. 1. Illustration of Support Vector Machine

$$\min_{w,b} \quad \frac{1}{2} w^\intercal w \tag{4a}$$
$$\text{s.t.} \quad y_i(w^\intercal x_i + b) \geq 1, \text{ for } i = 1, 2, \cdots, n \tag{4b}$$

where each tuple (x_i, y_i) is a data point from the training dataset, and n is the size of the training set.

C. *Linearly Non-separable Case*

The SVM classifier introduced in Section II-B is under the assumption that two classes are linearly separable. However, this would not always happen in the reality. For the case where two classes are not linearly separable, there are in general two solutions: (1) incorporating "soft margin", namely adding slack variables to tolerate possible classification errors; and (2) using the "kernel method", which transforms the linear boundary to a non-linear one. The following two subsections introduce each method in detail.

1) Soft Margin: If there does not exist a separating hyperplane between two classes, then Problem (4) is infeasible and thus incapable to find the optimal separating hyperplane in this case. Vapnik [24] proposed the idea of "soft margin" in 1992 to generalize SVM for the consideration of linearly non-separable case. In contrast, the linear separable case is often named "hard margin".

We introduce slack variables $\xi_i \geq 0$ to each data point (x_i, y_i) and relax Eqn. (4b) to be

$$y_i(w^\intercal x_i + b) \geq 1 - \xi_i, \text{ for } i = 1, 2, \cdots, n \tag{5}$$

ξ_i will be strictly positive when a misclassification happens, and $\xi_i = 0$ indicates a correct classification. To find the optimal separating hyperplane, we need to reduce the misclassification rate and maximize the margin. The objective of the hard margin case is generalized by adding a regularizer

Authorized licensed use limited to: Texas A M University. Downloaded on December 17,2021 at 04:45:21 UTC from IEEE Xplore. Restrictions apply.

$C\sum_{i=1}^{n}\xi_i$. Then the SVM classifier with soft margin is to solve the following problem:

$$\min_{w,b} \quad \frac{1}{2}w^{\intercal}w + C\sum_{i=1}^{n}\xi_i \tag{6a}$$
$$\text{s.t.} \quad y_i(w^{\intercal}x_i + b) \geq 1 - \xi_i, \text{ for } i = 1, 2, \cdots, n \tag{6b}$$
$$\xi_i \geq 0, \text{ for } i = 1, 2, \cdots, n \tag{6c}$$

where C is a scalar and smaller C typically implies softer margins.

2) Kernel Method: Kernel method is another simple yet powerful idea. It is worth mentioning that kernel method could not only be applied on SVM, but also on a large variety of problems, e.g. kernel PCA [25], kernel ridge regression [26] , spectral clustering [27]. To have a better understanding on the kernel method, we first write the Lagrangian dual of Problem (4):

$$\min_{a} \quad L = \frac{1}{2}w^{\intercal}w + \sum_{i=1}^{n}a_i[y_i(w^{\intercal}x_i + b) - 1] \tag{7a}$$
$$\text{s.t.} \quad a_i \geq 0, \text{ for } i = 1, 2, \cdots, n \tag{7b}$$

where a_i is the Lagrangian multiplier associated with Eqn. (6b). Using KKT conditions, we could obtain

$$\sum_{i=1}^{n}a_i y_i x_i = w \tag{8}$$
$$\sum_{i=1}^{n}a_i y_i = 0 \tag{9}$$

Then Problem (7) could be re-written as

$$\min_{a} \quad \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_i a_j y_i y_j (x_i^{\intercal}x_j) - \sum_{i=1}^{n}a_i \tag{10a}$$
$$\text{s.t.} \quad \sum_{i=1}^{n}a_i y_i = 0 \tag{10b}$$
$$a_i \geq 0, \text{ for } i = 1, 2, \cdots, n \tag{10c}$$

The key idea of the kernel method is to generalize the inner product $x_i^{\intercal}x_j$ in Eqn. (10a) to be $\Phi(x_i, x_j)$. Popular choices of kernel functions $\Phi(\alpha, \beta)$ include polynomial kernel $\Phi(\alpha, \beta) = (\alpha^{\intercal}\beta + \gamma)^d$, sigmoid kernel $\Phi(\alpha, \beta) = \tanh(k\alpha^{\intercal}\beta - \delta)$ and radial basis function (RBF) kernel (also called Gaussian kernel) $\Phi(\alpha, \beta) = \exp(-\gamma||\alpha - \beta||^2)$. The RBF kernel has fewer parameters than the polynomial and sigmoid kernels. This advantage of RBF kernel often leads to high probability of finding the optimal hyperplane and less numerical issues [9].

The kernel method could be applied on the linearly non-separable case as well. The derivation is almost identical with the linearly separable case.

III. The SVM-based Relay

Over-current relays operate according to threshold settings and depend on the assumptions on topology, generation levels and load levels. Starting from these assumptions, the fault current is calculated via circuit analysis. As discussed in Section I, network topology, generation level and load level always change with time in the reality. These could render the operation of over-current relays ineffective and lead to possible maloperations and misstrips. Motivated by this, we introduce a novel protection relay in this section. The proposed data-driven SVM-based relay (in short, the SVM-based relay) operates based on data-driven criteria and could outperform over-current relays.

A. Different Conditions for Protection Relays

For a protection relay, there are only two possibilities of a power system: (1) *normal condition*; and (2) *fault condition*. The *emergency condition* for a particular relay is a special case of the fault condition when the fault is happening on the same line as the relay locates. We term the non-emergency fault conditions and normal conditions as *other conditions*[1] for simplicity.

There are in general two classes of relays: main relays and backup relays. *Main relays* are designed to handle the emergency condition, and backup relays operate in case main relays failed to respond to failures in the neighboring lines. Ideally, a main relay should only operate during emergency condition to isolate the fault from the power system. And it should not operate under other conditions. As demonstrated in Section IV, the SVM-based relay could accurately distinguish the emergency condition from other conditions, however the over-current relays have difficulties and could maloperate or misstrip.

B. Design of the SVM-based Relay

The SVM-based relay is trained over a set of historical data, which contains both feature vectors and label vectors. The label vectors contain two classes: class 1 indicating emergency condition and -1 representing other conditions. The feature vectors refer to the active and reactive power measurements from one pair of relays on the same line. This pair of relays uses the SVM classifier model built by the training data measured locally to make operating decisions. More specifically, the feature vectors include the 3-phase measurements of both active and reactive power on the same line. Three-phase measurements are necessary because there might be single-phase or double faults happening. Both active and reactive power measurements are required since they both contain information of currents and voltages.

The propose SVM-based relay requires two types of communications (shown in Fig. 2): (1) between a relay and the central computer; and (2) within each pair of relays on the same line. When applied in a real distribution system, measurement units of the relays send the information to central computer, mainly including active power and reactive power of each phase. Based on the most recent measurement data, the central computer updates the SVM model for each relay and transmits the updated model to each relay. Since each SVM classifier is essentially a vector w and a scalar b, the communication

[1] Most protective relays (even backup relays) only need to identify the fault condition where faults occurring nearby, rather than identifying all fault conditions.

Authorized licensed use limited to: Texas A M University. Downloaded on December 17,2021 at 04:45:21 UTC from IEEE Xplore. Restrictions apply.

burden is not a concern in this case. Based on the SVM classifiers, each pair of relays make operation decisions by the real-time internal synchronous communications.

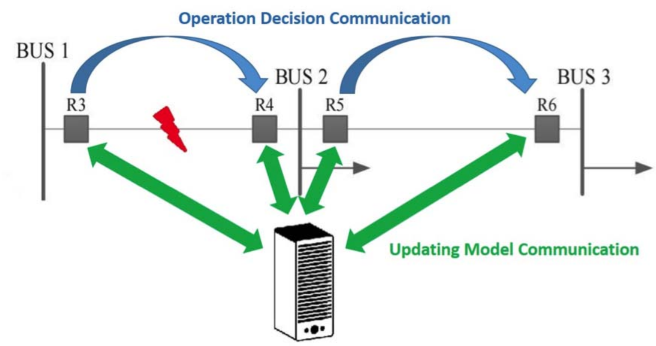

Fig. 2. Communication Infrastructure

Since the SVM model is updated using real-time measurements, which implicitly considers the impacts of changing system topology and DERs, the SVM-based relay is able to handle time-varying system conditions and make correct decisions. This proposed operating strategy is robust to possible communication failures between relays and the central computer. Indeed the communication burden could be further reduced by transmitting data only when large changes in the measurements are observed.

IV. Case Studies

This section presents case studies conducted on the IEEE 123-node Test Feeder (Fig. 3). Gridlab-D [28] is used to simulate the normal and fault conditions and collect data. We use the Statistics and Machine Learning Toolbox of Matlab to train the SVM classifier. Section IV-B-IV-D compare the performance of the SVM-based relay with over-current relays in three different aspects.

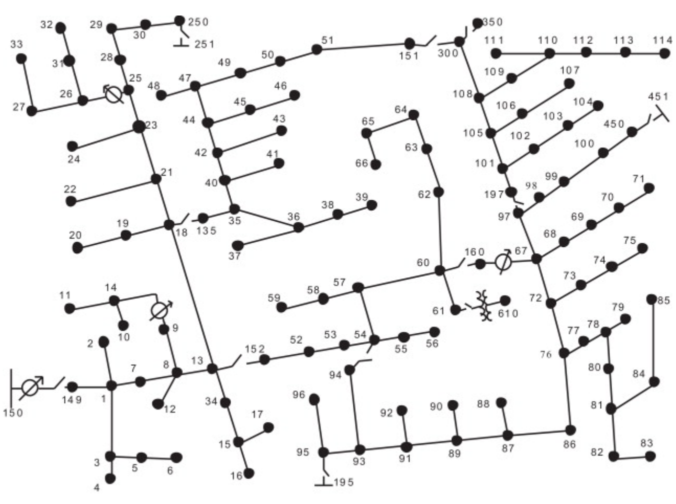

Fig. 3. IEEE 123-Node Test Feeder

A. Simulation Settings

For a thorough performance comparison between SVM-based relays with over-current relays, we simulate the IEEE 123-node Test Feeder with different settings:

1) Loads are uniformly distributed between 80% and 200% of the nominal load level. We assume loads are independent of each other and samples of loads are independent and identically distributed (i.i.d).
2) The outputs of DERs are also uniformly distributed from 80% to 120% of the pre-defined generation level. Samples of outputs of DERs are i.i.d. The penetration level of DERs is set to be 40%, i.e. the capacity of all DERs is equal to 40% of total load in the system.
3) Grounding faults can occur at different locations of different lines with the same probability. We assume it is unlikely that two faults happen at the same time.

We generate two *independent* datasets: training dataset and test dataset. 45 dB measurement noises are added to both the training and test dataset. The training dataset consists of 2000 points in the emergency condition and 228000 points in other conditions. The test dataset contains 500 points in the emergency condition and 56500 points of other conditions. The SVM classifier with RBF kernel is trained on the training dataset. The performance of both the SVM-based relay and over-current relay is evaluated on the test dataset. In this paper, we use the average accuracy of all relays in the system to detect internal faults as the main metric to quantify the performance. We use *maloperation* rates to denote the false positives, i.e. the relay believes there is a fault happening while it is actually not. We use *misstrip* rates to denote the false negatives, i.e. the relay believes no fault is happening while there indeed is one.

The threshold of conventional over-current relays is determined by the current value under the following conditions [29]: (1) fault occurs at the end of the line; (2) fault impedance is 20 ohms; and (3) load level is 120%, which is the assumed maximum loading condition for the relay. The over-current relays operate according to a simple logic with this calculated current value.

B. Impacts of Different Short-Circuit Impedances

Table I compares the performance with different fault impedances. Intuitively, high fault impedance would result in low fault current, which could be smaller than the threshold used in over-current relays. Because of this, high fault impedance might lead to frequent misstrip of the over-current relays. As shown in Table I, higher impedance values significantly increase the misstrip rates of over-current relays. The results in Table I also demonstrate the robustness of the SVM-based relays to various fault impedance via the comparison with the performance of the over-current relays.

The illustration of SVM classifier with the RBF kernel is shown in Fig. 4. For better illustration, we visualize the SVM classifier on a 2D plane (single-phase active power measurements from one pair of relays). The red dots represent other conditions, and green ones denote the emergency condition. The actual SVM classifier was trained in a 12-dimensional

Authorized licensed use limited to: Texas A M University. Downloaded on December 17,2021 at 04:45:21 UTC from IEEE Xplore. Restrictions apply.

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT RELAYS WITH DIFFERENT FAULT IMPEDANCES

	Over-current Relays		SVM-based Relays	
Fault Impedance (Ohm)	20	200	20	200
maloperation (%)	39.95	37.37	1.66	1.55
misstrip (%)	0.1	9.63	0	4.19

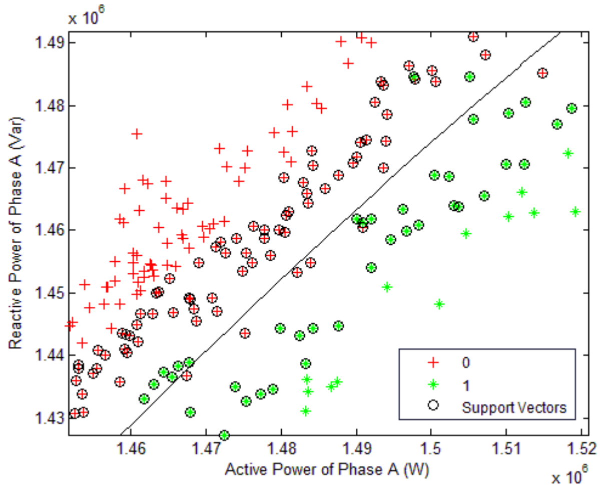

Fig. 4. Visualization of the SVM Classifier (RBF Kernel)

space, which consists of 3-phase active and reactive power measurements from one pair of relays. Fig. 4 shows a case where two classes are non-separable, a RBF kernel with soft margin is applied in this case.

C. Impacts of Load Levels

With an increasing amount of behind-meter PVs and electric vehicles in the distribution grid [30], the load levels could vary notably within a short period of time. It is necessary to verify the performance of the relays under various load levels. We define normal load level as 80% to 120% of the nominal load level; high load level refers to 120% to 200% of the nominal load level. Simulation results are shown in Table II. Conventional over-current relays have much higher maloperation and misstrip rates with load levels higher than 120%, which is the assumed load levels for threshold calculation. Results in Table II reveal the difficulties of over-current relays in distinguishing high currents caused by faults from high current due to high load levels. SVM-based relays are around 99% accurate with different load levels.

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT RELAYS WITH DIFFERENT LOAD LEVELS

	Over-current Relays		SVM-based Relays	
Load Levels	Normal	High	Normal	High
maloperation (%)	39.95	74.12	0.21	1.24
misstrip (%)	0	0.93	0	0

D. Impacts of DERs

Another main challenge in the distribution system is due to the increasing penetration of DERs [31], [32]. In modern distribution systems, the power power could be bidirectional instead of unidirectional in conventional distribution systems. This could cause maloperation and misstrip and thus pose new challenges for designing protection systems of distribution grids. Table III presents the simulation results in two scenarios: with and without DERs in the distribution system. Compared with SVM-based relays, over-current relays have much larger maloperation and misstrip rates. Increasing DERs is similar with the case of decreasing loads, where both could lead to low current in normal conditions. The pre-calculated thresholds of over-current relays might be inappropriate, which could lead to increasing misstrip rates. SVM-based relays update the operating rules in real time and the impacts of the penetration of DERs could be considered by updating classifier models.

TABLE III
PERFORMANCE COMPARISON OF DIFFERENT RELAYS WITH/WITHOUT DERS

	Over-current Relays		SVM-based Relays	
DERs	Without	With	Without	With
maloperation (%)	39.95	28.38	0.44	0.45
misstrip (%)	0	8.97	0	0

V. DISCUSSIONS

A. Pros and Cons

Compared with over-current relays, SVM-based relays have the following three advantages:

- The operating rule for the SVM-based relay is determined by the trained SVM model. It is purely data-driven and does not require any information on the network topology, load level, etc;
- The SVM-based relay makes better use of available measurement data. Although only two types of measurements (i.e. active and reactive power) are being used in this paper, the proposed method is generalizable to include more types of data (e.g. current, voltage) and its performance could be further improved.
- The SVM model can be updated in real-time by the streaming data to maintain high accuracy of operation.

The SVM-based relay proposed in this paper could be further improved in the following aspects:

- Operation decisions of SVM-based relays depend on the internal synchronous communications of two relays on the same line. This might increase the total costs of SVM-based relays. Furthermore, communication failure or asynchrony could increase the maloperation or misstrip rates.
- The SVM-based relay proposed in this paper is designed as main protection relays. Our future works include generalizing current design of SVM-based relays towards backup relays.

Authorized licensed use limited to: Texas A M University. Downloaded on December 17,2021 at 04:45:21 UTC from IEEE Xplore. Restrictions apply.

B. Distributed Operation of the SVM-based Relay

The proposed operating strategy above requires a central computer to update the SVM model based on streaming measurements. However, we want to emphasize that it is compatible with operating in a distributed manner. If the smart relay is programmed to solve a small-scale quadratic programming problem (which is efficiently solvable in polynomial time), then the central computer is not necessary.

VI. Conclusion

Due to increasingly complex system topologies and conditions, over-current relays (e.g. over-current relays) are facing great challenges. Frequent low fault currents happening in the distribution system migh lead to high possibilities of misstrip; changing load level could increase the risk of maloperation; the penetration of DERs influences the direction of power flow and could have impacts on both misstrip and maloperation. Poor performance of over-current relays are mainly due to prior assumptions, which might be inappropriate in modern distribution grids. In this paper, we propose a novel SVM-based relay which could make better use of the available streaming data in distribtion grids and outperform over-current relays. Case studies on the IEEE 123-node Test Feeder demonstrate the robustness of the propose SVM-based relays to altering system conditions such as load levels and fluctuating DERs.

References

[1] H. Wan, K. Li, and K. Wong, "An adaptive multiagent approach to protection relay coordination with distributed generators in industrial power distribution system," *IEEE Transactions on Industry Applications*, vol. 46, no. 5, pp. 2118–2124, 2010.

[2] V. Centeno, A. Phadke, A. Edris, J. Benton, M. Gaudi, and G. Michel, "An adaptive out-of-step relay [for power system protection]," *IEEE Transactions on Power Delivery*, vol. 12, no. 1, pp. 61–71, 1997.

[3] B. F. Wollenberg and T. Sakaguchi, "Artificial intelligence in power system operations," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1678–1685, 1987.

[4] H.-T. Yang, W.-Y. Chang, and C.-L. Huang, "A new neural networks approach to on-line fault section estimation using information of protective relays and circuit breakers," *IEEE Transactions on Power delivery*, vol. 9, no. 1, pp. 220–230, 1994.

[5] L. Montoya, D. Montenegro, and G. Ramos, "Adaptive protection testbed using real time and hardware-in-the-loop simulation," in *PowerTech (POWERTECH), 2013 IEEE Grenoble*. IEEE, 2013, pp. 1–4.

[6] W. El-Khattam and T. S. Sidhu, "Restoration of directional overcurrent relay coordination in distributed generation systems utilizing fault current limiter," *IEEE Transactions on Power Delivery*, vol. 23, no. 2, pp. 576–585, 2008.

[7] H. Zhan, C. Wang, Y. Wang, X. Yang, X. Zhang, C. Wu, and Y. Chen, "Relay protection coordination integrated optimal placement and sizing of distributed generation sources in distribution networks," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 55–65, 2016.

[8] M. H. U. Ahmed, M. G. R. Alam, R. Kamal, C. S. Hong, and S. Lee, "Smart grid cooperative communication with smart relay," *Journal of Communications and Networks*, vol. 14, no. 6, pp. 640–652, 2012.

[9] Y. Zhang, M. D. Ilic, and O. K. Tonguz, "Mitigating blackouts via smart relays: a machine learning approach," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 94–118, 2011.

[10] K. Seethalekshmi, S. Singh, and S. Srivastava, "A classification approach using support vector machines to prevent distance relay maloperation under power swing and voltage instability," *IEEE Transactions on Power Delivery*, vol. 27, no. 3, pp. 1124–1133, 2012.

[11] R. Salat and S. Osowski, "Accurate fault location in the power transmission line using support vector machine approach," *IEEE Transactions on power systems*, vol. 19, no. 2, pp. 979–986, 2004.

[12] Y. Zhang, M. D. Ilic, and O. Tonguz, "Application of support vector machine classification to enhanced protection relay logic in electric power grids," in *Power Engineering, 2007 Large Engineering Systems Conference on*. IEEE, 2007, pp. 31–38.

[13] B. Ravikumar, D. Thukaram, and H. Khincha, "Comparison of multiclass SVM classification methods to use in a supportive system for distance relay coordination," *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1296–1305, 2010.

[14] H. Jiang, J. J. Zhang, W. Gao, and Z. Wu, "Fault detection, identification, and location in smart grid based on data-driven computational methods," *IEEE Transactions on Smart Grid*, vol. 5, no. 6, pp. 2947–2956, 2014.

[15] H. Jiang, X. Dai, D. W. Gao, J. J. Zhang, Y. Zhang, and E. Muljadi, "Spatial-temporal synchrophasor data characterization and analytics in smart grid fault detection, identification, and impact causal analysis," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2525–2536, 2016.

[16] A. Phadke, B. Pickett, M. Adamiak, M. Begovic, G. Benmouyal, R. Burnett, T. Cease, J. Goossens, D. Hansen, M. Kezunovic, and others, "Synchronized sampling and phasor measurements for relaying and control," *IEEE Transactions on Power Delivery*, vol. 9, no. 1, pp. 442–452, 1994.

[17] M. Tasdighi and M. Kezunovic, "Preventing transmission distance relays maloperation under unintended bulk DG tripping using SVM-based approach," *Electric Power Systems Research*, vol. 142, pp. 258–267, 2017.

[18] P. Dash, S. Samantaray, and G. Panda, "Fault classification and section identification of an advanced series-compensated transmission line using support vector machine," *IEEE transactions on power delivery*, vol. 22, no. 1, pp. 67–73, 2007.

[19] X. Geng and L. Xie, "A data-driven approach to identifying system pattern regions in market operations," in *2015 IEEE Power & Energy Society General Meeting*. IEEE, 2015, pp. 1–5.

[20] ——, "Learning the LMP-Load Coupling From Data: A Support Vector Machine Based Approach," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1127 – 1138, 2017. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7478156/

[21] J. Shi, W.-J. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," *IEEE Transactions on Industry Applications*, vol. 48, no. 3, pp. 1064–1069, 2012.

[22] H. Jiang, Y. Zhang, E. Muljadi, J. Zhang, and W. Gao, "A short-term and high-resolution distribution system load forecasting approach using support vector regression with hybrid parameters optimization," *IEEE Transactions on Smart Grid*, 2017.

[23] M. Kezunovic, L. Xie, and S. Grijalva, "The role of big data in improving power system operation and protection," in *Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid (IREP), 2013 IREP Symposium*. IEEE, 2013, pp. 1–9.

[24] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 1992, pp. 144–152.

[25] P. R. Gribik, W. W. Hogan, and S. L. Pope, *Market-clearing electricity prices and energy uplift*. December, 2007.

[26] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.

[27] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

[28] D. P. Chassin, K. Schneider, and C. Gerkensmeyer, "GridLAB-D: An open-source power systems modeling and simulation environment," in *Transmission and distribution conference and exposition, 2008. t&d. IEEE/PES*. IEEE, 2008, pp. 1–5.

[29] J. He, X. Dong, Y. Li, and B. Li, *Principles of Power System Protection*. China Electric Power Press, 2010.

[30] Q. Yan, C. Qian, B. Zhang, and M. Kezunovic, "Statistical analysis and modeling of plug-in electric vehicle charging demand in distribution systems," in *Intelligent System Application to Power Systems (ISAP), 2017 19th International Conference on*. IEEE, 2017, pp. 1–6.

[31] P. Dehghanian, M. Fotuhi-Firuzabad, F. Aminifar, and R. Billinton, "A comprehensive scheme for reliability centered maintenance in power distribution systemsPart I: Methodology," *IEEE Transactions on Power Delivery*, vol. 28, no. 2, pp. 761–770, 2013.

[32] ——, "A comprehensive scheme for reliability-centered maintenance in power distribution systemsPart II: Numerical analysis," *IEEE Transactions on Power Delivery*, vol. 28, no. 2, pp. 771–778, 2013.

Authorized licensed use limited to: Texas A M University. Downloaded on December 17,2021 at 04:45:21 UTC from IEEE Xplore. Restrictions apply.