Electric Power Systems Research 142 (2017) 258–267

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Preventing transmission distance relays maloperation under unintended bulk DG tripping using SVM-based approach

Mohammad Tasdighi *, Mladen Kezunovic

Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA

A R T I C L E I N F O

Article history:
Received 23 May 2016
Received in revised form
19 September 2016
Accepted 21 September 2016

Keywords:
Bulk DG tripping
Phasor measurement unit (PMU)
Local measurements
Wide-area (WA) measurements
Support vector machine
DG point of common coupling (PCC)

A B S T R A C T

With high penetration of distributed generation (DG) into distribution systems, an unintended bulk DG tripping as a consequence of severe disturbances such as 3-phase faults in the transmission system is a concern since it can further cause unintended operation of transmission distance relays. To address this matter, this paper presents a novel protection scheme based on support vector machine (SVM) approach. The proposed scheme detects bulk DG tripping following a fault in the power transmission system, and then makes sure there is no follow on distance relay maloperation. It is also able to detect a fault if it happens during the blocking period and hence unblock the relay operation correspondingly. Wide-area (WA) measurements obtained from phasor measurement units (PMUs) are used, in addition to local measurements, to improve the scheme selectivity. The New-England 39 bus system is used to test the proposed scheme. Simulation results are discussed and illustrated.

1. Introduction

In today's modern power systems, DGs are growing rapidly based on economic and environmental incentives [1]. They are required to follow standards for connecting to the grid and have control and protection measures on their interconnections to be able to disconnect from the distribution grid in case of an inadvertent islanding [2–7]. Inadvertent islanding is called to the situation when DG continues energizing a portion of the system, e.g. the feeder that it is connected to, while being disconnected from the main grid [8]. The duration and probability of an inadvertent island occurrence must be minimized for several reasons such as mitigating power quality, maintaining protection settings, addressing auto reclosing issues, and most importantly ensuring the staff safety [8]. Several anti-islanding protection schemes which are mainly categorized into communication based and local measurement based methods have been proposed and developed based on this necessity [8,9]. Since deploying communication based methods, known as transfer trip, is not cost effective for widespread use, the local measurement based methods are commonly used for anti-islanding protection purposes at the distribution level [8]. Generally, the local measurement based methods are divided into active and passive ones for which the set of protection schemes consist of under and over frequency and voltage relays [8,9].

On the other hand, sensitive protection schemes being in charge of tripping DGs could act as a threat to the upstream network's post disturbance response as the DG penetration grows in the system [10–14]. This is because of their probable maloperation as a consequence of a severe disturbance happening upstream which could trigger unintended bulk DG tripping on the distribution side and impose extra stress on the system. This is one of today's immediate concerns of some independent system operators (ISOs) [13,14]. For their network, the DGs are connected to the distribution grid according to IEEE 1547 [2] which does not allow DG's ride-through during voltage/frequency deviations for the previously mentioned reasons, especially ensuring the utility personnel's safety. They are studying the impacts of such events on their bulk power systems and trying to alleviate the corresponding consequences [13,14].

One critical consequence of the additional imposed load flow stress on the upstream network as a result of the unintended bulk DG tripping is the probable unforeseen interference with the conventional distance protection [15–17] as it will be discussed in more detail in Section 2. It is worth pointing out that the above mentioned concern does not extend to conventional generator tripping in the system for two main reasons. Firstly, according to the NERC standard, conventional generators are required to stay connected to the grid throughout almost all the disturbances to maintain the system's synchronization by their turbine-generator inertia [18,19].

* Corresponding author at: Wisenbaker Engineering Research Center, Texas A&M University, College Station, Texas, USA.
E-mail addresses: m.tasdighi@tamu.edu, mohammad.tasdighi@gmail.com (M. Tasdighi), kezunov@ece.tamu.edu (M. Kezunovic).

http://dx.doi.org/10.1016/j.epsr.2016.09.024
0378-7796/

They participate in load frequency control (LFC) and automatic governor control (AGC) actions performed by ISOs sending control signals and set points to the generators in real-time to set their outputs [20]. Secondly, the dynamic planning studies performed by ISOs according to the NERC standard [21] already check the dynamic behavior of the system to be reliable and safe for N-1 contingency cases including each conventional generator tripping. Before a conventional generator is connected and added to the grid, it will be verified that its unintended tripping will not lead to any system instability or cascade event and the required precautions and corrective actions would be planned [21]. The distance protection on transmission side is coordinated for these N-1 contingency cases [22]. However, planning and protection studies for transmission network are based on the network models that do not contain DG protection models, and detailed protection information is not included in the bulk DG planning studies [23,24].

According to what is mentioned thus far, it could be concluded that it is necessary to make sure the dependability and security of the protection on the transmission side is not affected by such unintended events to prevent damage extension from distribution to the bulk power system. Although the impacts of unintended DG tripping on transmission protection coordination has been brought up in the literature [15–17], no protection scheme has been specifically proposed against undesirable tripping of distance relays under such circumstances. In this study, a novel SVM-based scheme is proposed to maintain the transmission protection security and dependability under unintended bulk DG tripping on the distribution side, which may occur as a result of maloperation of the deployed anti-islanding schemes.

SVM performance, when compared to the other conventional classifiers such as neural networks, fuzzy logic, etc., the performance of which might suffer from handling huge feature spaces, is not significantly affected by classified vectors dimension [25]. Neural-network approaches have been shown to be effective in many applications; however, their main disadvantage is the need for significant training burden (data and time) for a reliable performance of the approach especially when the operating conditions vary widely [25]. Furthermore, the great advantage of SVM, which makes it more powerful than other traditional methods based on risk minimization is that it deploys various ideas such as the Vapnik-Chervonenkis theory, statistical learning, maximum margin optimal hyperplane, kernel functions and so on [25].

SVM has been employed as a supervised learning method for different power system protection purposes recently [26–30]. Ravikumar et al. [26] have used SVMs for coordination of distance relays in the transmission system. Samples of apparent impedance seen by the relay during faults are used as SVM input data. The same authors have evaluated and compared various methods of implementing multiclass SVMs in studying the coordination of distance relay settings [27]. Some other studies have used SVM technique for improving protection of transmission lines compensated by series capacitors [28,29]. The authors in [28] have proposed a method in which fault location is identified on the compensated line. In [29], authors have presented a combined wavelet-SVM technique which uses three line current samples to detect the faulted zone on a series compensated transmission line. Seethalekshmi et al. [30] have deployed SVM technique to improve distance relay power swing and voltage instability detection.

SVM technique is deployed here to implement a protection scheme that enables the vulnerable distance relays (target relays), the backup settings (second or third zones) of which might get affected under unintended DG tripping events, to distinguish such events from faults and block/un-block the relay operation correspondingly. A recently proposed and implemented novel setting coordination check module [31,32] is used to identify the target relays in the test system. Selective WA measurements obtained by PMUs in addition to local measurements at the distance relay location are used to improve the proposed scheme accuracy. The scheme's robustness against PMU data loss or unavailability as well as cost-wise use of WA measurement technology has been taken into consideration in the proposed method. The SVM is trained such that it distinguishes the faults from the DG tripping cases and acts as the supervisory control of the distance backup protection. In the case of unintended DG tripping interference with the distance relay setting coordination, the proposed scheme blocks the conventional trip signal resulting from the distance mho elements' pickup and prevents any follow on the distance relay maloperation. Furthermore, unlike conventional blocking schemes, the proposed method is able not only to block the relay operation due to DG tripping interference, but also to detect a fault during the blocking period and unblock the relay correspondingly. The proposed scheme is easily and quickly trainable for various possible scenarios of system operation in practice and gives significant selectivity. Furthermore, it could be considered as another complementary application of SVM along with previously proposed ones to obtain a comprehensive supervisory control protection scheme and improve the protection security and dependability.

The rest of the paper is organized as follows. Section 2 provides a detailed problem description of how the protection coordination of the upstream transmission network may be affected by unintended bulk DG tripping on the distribution side. Section 3 presents a brief introduction to the SVM technique. The proposed SVM-based protection scheme and the process of identifying the target relays in the system are discussed in Section 4. Section 5 presents the simulation results of implementing the proposed scheme on the New-England test system. Concluding remarks and the paper main contributions are summarized in Section 6. References are given at the end.

2. Problem description

The anti-islanding protection schemes are responsible for detaching the DGs from the grid in case of an inadvertent islanding. The basic idea is sensing the voltage and frequency deviations and checking them against the threshold values to come up with the control action. The critical need to prevent islanding occurrence, especially in order to guarantee the personnel's safety, along with some probable hard-to-detect cases of islanding [17] drives the anti-islanding protection control and measures to be sensitive enough to detect the islanding cases. On the other hand, these sensitive protection measures could affect the DG output unnecessarily under certain circumstances and aggravate the power system dynamic behavior during or after disturbances. Under frequency and voltage sensitivities are two important indicators of such conditions. The former corresponds to a generation-load mismatch situation which may trigger bulk DG tripping, which deteriorates the situation further. Such cases of unintended DG tripping could be mitigated by taking proper immediate load shedding actions, which is not the focus of this study.

The voltage sag caused by severe disturbances such as 3-phase faults at the transmission side could propagate to the distribution level and interfere with DG's under-voltage protection measures, which may lead to unintended DG tripping. This might not raise any significant issue if the existing DG in the system is of small scale and the system is well-designed to handle that. However, in case of high penetration of DG in the distribution network, connected to upstream through a point of common coupling (PCC), the large scale tripping of the DG units puts an extra power flow burden on the transmission lines. As a result, protection coordination of distance relays' backup protection zones on transmission side might get affected. The sudden power flow increase to compensate

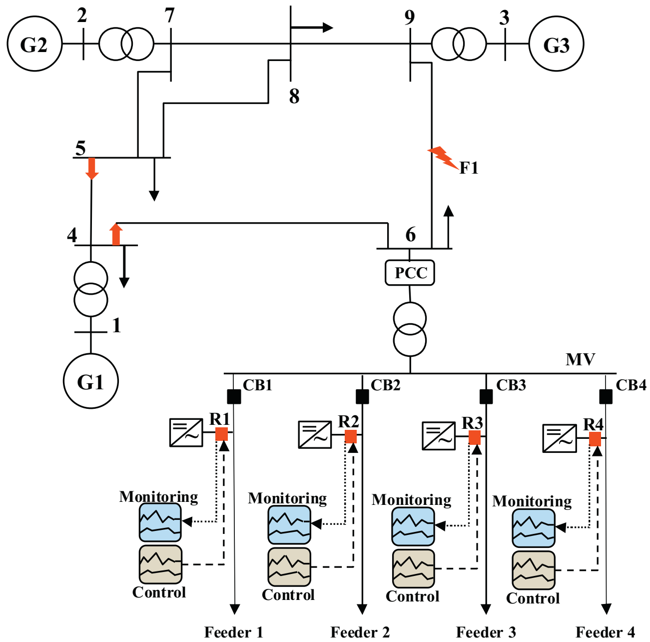

Fig. 1. Possible scenario of unintended bulk DG tripping as a consequence of under voltage trip sensitivity.

the lack of DG in the system which is already under stress from previous disturbance could initiate distance relay miss-operation and lead to cascade events. Other disturbances such as major switching actions (lines or generators tripping) could also lead to significant voltage deviations which might be potential cause of DG tripping.

Fig. 1 helps to illustrate the problem under consideration. Anti-islanding protection scheme makes sure that the DG connected to a feeder (Feeder 1-Feeder 4) would trip if the feeder's source-side circuit breaker (CB1-CB4) opens, usually as a result of a fault on the feeder. A short-circuit happens on the line 9-6 and it is tripped to clear the fault. The voltage drop and deviations as a result of fault occurrence and clearing event propagates to the distribution side. Assuming the DGs are tripped mistakenly by their anti-islanding protection systems, a sudden power flow increase is imposed on the lines 5-4 and 4–6 to compensate for the lack of DGs while the system is still under the stress of the previous disturbance. This might cause an interference with setting coordination of distance relays on these lines (marked by red arrows) as a result of unexpected dynamic change of the impedance trajectory and trigger their miss-operation, isolation of buses 4 and 6, and lead to the total system collapse consequently. It should be noted that, this is just a simple graphical example to help picturing the problem tentatively; of course, various parameters including the dynamic behavior of the system, impedances of the lines, the settings of the distance relays, the capacity and instant of the tripped DG, loadability of the lines, etc. are important in determining whether it would cause the distance relays missoperation or not. A real demonstration of this scenario on New-England 39-bus system will be presented in Section 5.

The problem described above highlights the necessity to manage the protection on transmission side to be able to come into the action and act quickly in case of an unintended operation by anti-islanding schemes on the distribution side to save the upstream network. It should be able to distinguish such cases from faults and block/unblock tripping signals of vulnerable relays' backup protective zones accordingly.

3. Support vector machine (SVM) technique

3.1. Brief overview

SVM is a relatively new and promising machine learning technique to be deployed as a pattern recognition and classification tool. It is based on the statistical learning theory for 'distribution-free learning from data' proposed by Vapnik [33]. In this method, first, the input data is mapped into feature space which is a high-dimensional dot product space and then it is classified through a hyper-plane. Using optimization theory, the maximum separation is obtained by the optimal hyper-plane.

Suppose $x_i \in R^n$ and $i \in \{1, \ldots, l\}$ is the input data including l data points which could be classified into two classes, class I and class II, with the labels of $y_i = 1$, and $y_i = -1$. The goal of SVM linear separation is to identify the optimal hyper-plane which creates the maximum separation between the data points in regards to their classes. For the above mentioned classes, such a separating hyper-plane could be achieved by finding out proper values for w, vector of weights, and b, biased scalar, in the following equation:

$$f(x) = w^T x + b = 0 \tag{1}$$

For a separating hyper-plane:

$$\begin{cases} f(x_i) \geq 1 & \text{if } y_i = +1 \\ f(x_i) \leq -1 & \text{if } y_i = -1 \end{cases} \tag{2}$$

Therefore, $y_i f(x_i) = y_i \left(w^T x_i + b\right) \geq 1$ for $i = 1, \ldots, l$. From the geometry, it is found that: $m = 2\|w\|^{-1}$ in which m represents the separation margin. So, maximizing m which means better generalization capability of SVM requires to minimize $\|w\|$. Hence, finding the optimal hyper-plane could be formulated as the following convex optimization problem:

$$\begin{aligned} &\min \quad \frac{1}{2} w^T w \\ &s.t. \quad y_i(w^T x_i + b) \geq 1 \quad \forall i \end{aligned} \tag{3}$$

There exist no hyper-plane if it is not possible to separate data linearly; i.e., the constraints in (2) cannot be satisfied all together. In such cases, a penalty factor C and slack variables ξ_i are deployed to introduce a soft margin. The optimization problem then changes to:

$$\begin{aligned} &\min \quad \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i \\ &s.t. \quad y_i(w^T x_i + b) \geq 1 - \xi_i \quad \text{for} \quad i = 1, \ldots, l \\ &\xi_i \geq 0 \text{ for } \quad i = 1, \ldots, l \end{aligned} \tag{4}$$

In (2), ξ_i are non-negative variables which bring training errors into the scene. The penalty factor C, also called regularization factor, is always positive. In case it is small, the separating hyper-plane is more focused on maximizing the margin (m) while the number of misclassified points is minimized for larger C values. Support vectors which include the points closest to the optimal hyper-plane maintaining maximum margin, satisfying (2) with equality sign, are required to obtain the separating hyper-plane.

The classification problems in practice are usually not linear. To implement SVMs for such cases, so called kernel functions are deployed for mapping training data by the use of nonlinear transform function $\phi(x)$:

$$\phi(x_i) = (\phi_1(x_i), \ldots, \phi_m(x_i)), \text{ where } \quad m > n \tag{5}$$

The equation which could define a kernel function is: $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$. Having done such a mapping, the goal is to be able to implement the linear classification of the original input data x in the higher-dimensional space by the use of linear SVM formulations.

Although SVMs are designed to be deployed for the binary classifications, they could be used for multiclass classification purposes too. Generally, there are three approaches to implement a multiclass SVM: one-against-one (OAO), one-against-all (OAA), and one-step methods. The first two approaches are based on combining several binary SVMs; however, in the one-step method the SVM is designed in a way to include all the classes at once during the learning algorithm and solve only one optimization problem [33–35]. The performance comparison between these three methods has shown that the one-step approach gives better accuracy in addition to be faster than the others [27]. Hence, this method is chosen here.

In one-step method, the idea is to create p two-class rules which are separated by p decision functions. For example, the vectors of class k are separated from the other vectors by the k^{th} function $w_k^T\phi(x)+b$. However, all the decision functions are obtained by solving one problem as follows:

$$\begin{aligned} \min \quad & \frac{1}{2}\sum_{k=1}^{p} w_k^T w_k + C\sum_{i=1}^{l}\sum_{k\neq y_i}\xi_i^k w_{y_i}^T\phi(x_i) \\ & +b_{y_i} \geq w_k^T\phi(x_i)+b_m+2-\xi_i^k \\ s.t. \quad & \xi_i^k \geq 0 \text{ for } i=1,\ldots,l \,\&\, k\in\{1,\ldots,p\}\backslash y_i \end{aligned} \tag{6}$$

And the decision function is:

$$\text{argmax}_{k=1,\ldots,p}\left(w_k^T\phi(x)+b_m\right)$$

3.2. Kernel function selection

Various kernel functions have been proposed by researchers such as linear, polynomial, radial basis function (RBF), and sigmoid kernel functions. In this study, RBF kernel $F(x_i,x_j)=\exp\left(-\gamma\|x_i-x_j\|^2\right)$ for $\gamma>0$ is considered as a reasonable first choice because of several reasons. Deploying RBF kernel provides non-linear mapping of input data sets and is able to deal with the non-linear correlation of the class labels and features so it overweighs the linear kernel [35]. Besides, a linear kernel is considered as a subset of RBF because for a definite penalty factor, C', it could be represented as the RBF kernel having specific parameters (C,γ) [36]. Sigmoid kernel also performs like RBF for certain parameters [37]. Moreover, there are some parameters for which the sigmoid kernel is not the dot product of two vectors so it is not valid [33]. Polynomial kernel has more unknown parameters to be determined compared to RBF kernel and this makes the model selection for polynomial kernel more complex. Furthermore, polynomial kernel values might be not properly bounded. Last but not least, numerical difficulties for the RBF kernel are fewer than the others [35].

3.3. Parameter selection

C and γ are two unknown parameters which should be determined when using RBF kernel. Proper parameter search must be conducted on the grid of data to find the best of these values for a given problem. The focus is on finding (C,γ) values for SVM classifier to be able to predict the unknown data, i.e. testing data set, accurately. A common approach is to provide two sets of data which are called training and testing or known and unknown data sets respectively. The SVM performance is better evaluated by prediction accuracy obtained from classifying the unknown independent data set. This process is called cross-validation in its advanced form.

In this study, C and γ are obtained by conducting a grid-search using cross-validation. To implement n-fold cross-validation, the training data set is divided equally into n subsets of data. Then, to test each subset, the SVM is trained on the remaining ones (n-1 subsets) so each training instance is tested once and training accuracy represents the number of subsets which were classified correctly. This technique is useful in preventing the over-fitting problem [35].

4. Proposed scheme

4.1. Identifying the vulnerable relays

To implement the proposed protection scheme, first, the relays which settings coordination might get affected due to unintended DG tripping should be identified according to the network topology. In [31,32], an automatic distance setting coordination check module is proposed and implemented, and its performance is verified by comparison with a commercial package (CAPE) [22]. The module has been run on the real-sized networks as well as New-England 39-bus system. The module is able to identify the affected relay settings coordination issues following a network topology change such as generation trip, line switching, etc. In such studies, usually the relay settings are calculated based on the line ohms only. In practice as well as in the developed module, the short-circuit calculations and apparent impedances are deployed in calculating the backup protective zones (zones 2 and 3) settings of the relay. The network operating conditions such as power flow are also considered in the setting coordination check process.

The relays are assumed being set in forward direction and the backup zones settings are time-delayed, i.e. 20 and 60 cycles for zones 2 and 3 operation respectively [22]. With the use of the proposed setting coordination check module in [31,32], a list of relays vulnerable to unintended DG tripping based on the network topology and DG placement is determined. The relays with a change beyond 5% in their zone 2 or 3 settings are identified as critical relays and sorted correspondingly [31,32]. Then, the proposed protection scheme could be implemented to those critical relays.

4.2. SVM-based scheme

In this paper a SVM based protection scheme which enables the distance relay to distinguish between a fault and a DG tripping scenario when interfering with the protection coordination of distance backup protective zones is proposed. The detection is based on the DG tripping impact on the system dynamic behavior. As shown in Fig. 2, two multiclass SVMs are deployed, one is trained based on local data only (SVM-1) and the other one is provided with WA data as well (SVM-2). Based on whether the PMU data is being received at the relay location or not, the method could switch between the employed SVMs outputs through the multiplexer shown in Fig. 2. This is for maintaining the scheme's robustness under probable PMU data unavailability or loss; however, the accuracy may decrease to some extent when using local data only as will be discussed in Section 5. SVM-1 and SVM-2 are trained to classify fault, DG tripping, and other cases as "1", "0", and "-1" respectively. The outputs of these SVMs are filtered by a comparator as class label −1 is not of interest. The logical AND of the backup protective zones pickup signal and the output of the comparator, as shown in Fig. 2, determines the trip/block signal value, i.e. 1 or 0.

A proper modeling of the DG units is important to get a fair observation of their impact on the dynamic behavior of the network following a disturbance. In this study, the focus is on PVs in the distribution level (residential PVs) which are modeled as constant current loads corresponding to the negative power injections, which is used in other studies of this type [16,19,38]. The equiva-

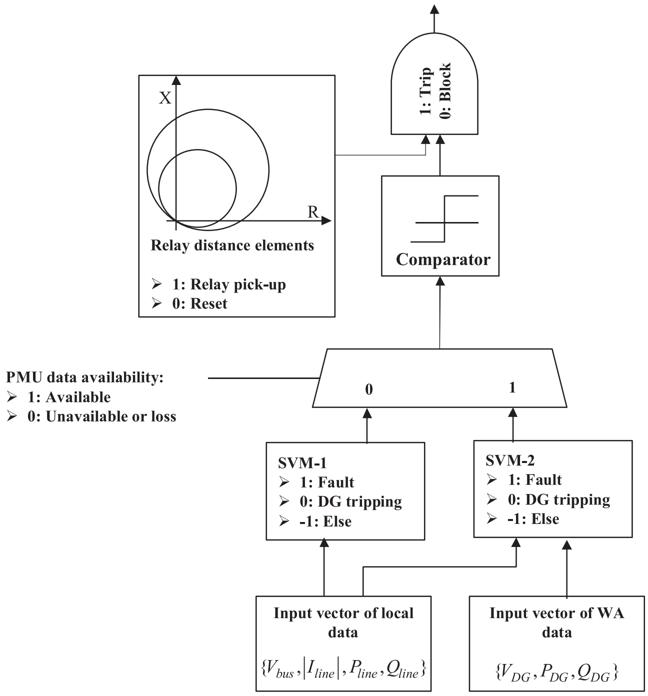

Fig. 2. Block diagram of the proposed scheme.

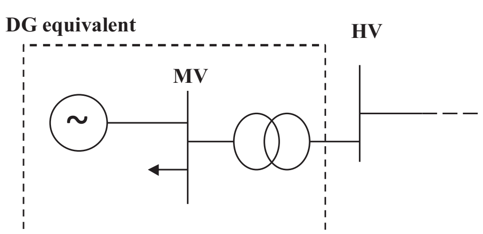

Fig. 3. DG plants equivalent from the transmission side.

lent of DG units aggregated based on their generation type from the transmission perspective could be represented as shown in Fig. 3 [16,19]. For studies of this type, the downstream distribution network, regardless of its connections, is modeled as the aggregated load and distributed generation imposed on the upstream network from the transmission point of view [16,19]. This accepted type of modeling is especially appropriate when the distribution grid is connected to a well interconnected and stable upstream network, as is the case in this study. The cluster of PVs is modeled as an equivalent power output equal to the sum of individual outputs of each one of the units [16,19]. It is worth to note that DGs are usually operated in constant power/power factor control mode [17]. From the transmission point of view, different types of DG would not be experienced significantly different from each other. Under specific cases of DG operation, if the majority of the DGs are facilitated with voltage regulation or speed controls based on their type, they should be modeled correspondingly [16]. We have focused on PVs because of their modeling simplicity. They are considered as the most promising type of DGs growing fast in the distribution level because of their economic and environmental incentives [19].

Local measurements and calculations based on them at the relay point are the required elements of almost all of the protection schemes. Features selected from the local measurements as inputs for the SVM-1 and SVM-2 are: V_{bus},$|I_{line}|$,P_{line}, and Q_{line} representing the bus voltage phasor, line current phasor magnitude, line active and reactive power flow respectively. Thanks to the PMU technology, WA measurements from various points of the system could be provided in today's power system operation. When employing WA measurements technology, implementation cost must be considered for the method to be economically justifiable. In other words, the more PMUs are deployed; the significantly higher implementation cost would be experienced although a better system behavior observation may be obtained. It is assumed that phasor measurements from PCC and target relay bus in regards to the reference bus are available and by PCC, as shown in Fig. 1, only the distribution to upstream connection bus is meant rather than all of the DG units' interconnections individually. Therefore, implementing the proposed method would be economically practical. Net active (P_{DG}) and reactive (Q_{DG}) power injections from the PCC into the transmission grid calculated from the PCC's PMU measurements are two good features to be utilized to improve the SVM-2 pattern recognition and classification accuracy. The other proper feature is the voltage phasor at the PCC on the grid side (V_{DG}). Deploying these measurements and calculations is specifically beneficial to improve the SVM's performance accuracy when classifying under more complicated scenarios such as detecting a second fault when the system is already under the stress of a post fault and subsequent DG tripping events. As it will be shown in Section 5, the proposed scheme is able to detect such cases and unblock the trip signal so the protection security and dependability is well maintained.

Depending on the application of the PMU, the role of communication requirements and latencies could get highlighted. The delay related to PMU deployment is caused from three main processes: phasor creation, transmission of data through the available communication link, and merging of data streams in phasor data concentrators (PDCs) [39]. The PMUs use fast mathematical algorithms, such as discrete furrier transform (DFT) and calculate the voltage and current phasors from RMS measurements obtained by voltage and current instrument transformers [39]. Then the phasor measurements are transmitted according to IEEE C37.118 [40] data format to PDCs via available communication link and the delay depends on the link's data transfer capability, the size of the PMU data output, as well as physical distance between PMU and PDC. The PDC delay, at the target substation in this study, corresponds to implementing time tag on the data and preparing a system-wide measurement [41]. There are various communication options available for wide area measurement system (WAMS) including telephone lines, fiber-optic cables, satellites, power lines, and microwave links [39]. Studies show that the average combined delay caused by the above mentioned reasons over even long transmitting distances (in the order of 1000 miles) when using a communication media with a band-width of 56 Kbps (data rate in telephone lines) is around 5–7 cycles of a 60 Hz system [41]. Therefore, deploying wide area measurements in the proposed method is a proper fit with regards to the method's application for the purpose of improving distance relay back-up protection which, as mentioned before, operates with a time delay (20 and 60 cycles for zones 2 and 3 respectively). Deploying advanced communication media such as fiber-optic cable by utilities provides a data transfer speed up to 2Mbps and significantly reduces the delay by removing the delay corresponding to the PMU data size [39].

The SVMs training scenarios includes different DG tripped capacities following 3-phase faults on the transmission side at various points in the vicinity of the DG placement in order to have realistic scenarios of the severe disturbance impact propagation from transmission to distribution level. The possible DG tripping instant following the disturbance varies in a range assumed according to the standards for anti-islanding protection schemes. Having prepared the training data set the SVMs go through the learning process and their performances are verified on the testing data set as will be discussed in the following Section.

Employing SVM technique to approach the problem under consideration from the distribution side, when using local data only,

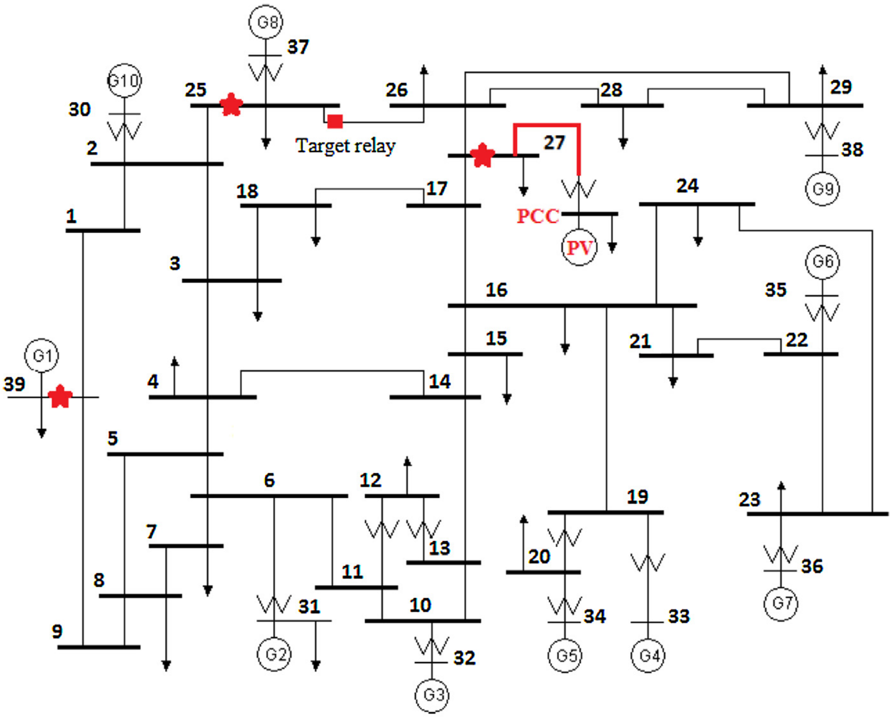

Fig. 4. One-line diagram of New-England 39 bus system with DG penetration.

Table 1
Vulnerable relays to DG tripping.

Rank	Critical Relay
1	R_{25-26}
2	R_{29-26}
3	R_{16-17}

may not be a proper fit. To maintain its operation accuracy under the probable upstream disturbances might be challenging. After all, the primary issue was raised when deviations propagating from transmission to the distribution side are close to and almost not differentiable from those caused by islanding situation which is a probable thread to any anti-islanding protection scheme. Under such circumstances, the accuracy of SVM may be affected if only relying on local measurements because training the SVM to differentiate between such cases actually means training it for instances with similar features yet different labels which lowers the classification accuracy. The unintended DG tripping risk still remains unless remote measurements are provided for each DG unit's interconnection relay which is the same as the costly method of transfer trip.

5. Case study

The simulations have been conducted on the New-England 39-bus test system [42], Fig. 4. Having conducted a sensitivity analysis on the test system using the setting coordination check module from [31,32], the buses for clustered DG location with higher impact on the network distance relay settings and their corresponding list of critical relays (target relays) are identified. It was concluded that clustered PVs on bus 27, as shown in Fig. 4, is one of the locations with highest impact on distance relay settings for unintended PV tripping cases and the corresponding list of critical relays to this location is brought in Table 1.

Maximum PV penetration in the system is assumed to be a considerable amount of 500 MW all of which has been tripped to sort the vulnerable relays in the system as the worst case scenario. The penetration level could be defined in various ways based on the system's total generation, system's peak load, or amount of energy served [19]. For example, considering the system's total generation, the penetration level is obtained from the following equation:

$$\text{DG Penetration}(\%) = \frac{\text{Total DG generation (MW)}}{\text{Total generation (MW)}} \tag{7}$$

That is around%10 in this study.

For the purpose of this study, the amount of the DG tripped capacity, the instant of the tripping following the disturbance on transmission side, current distance relay settings, etc., which play the key role in the protection coordination interference as it would be illustrated in Section 5.1., are more important than the total level of penetration.

The proposed method is implemented for the most critical relay (R_{25-26}), which is considered as the target relay. As mentioned before, the PMUs are assumed to be located on the reference bus (bus 39), the target relay bus (bus 25) and DG PCC as shown by stars in Fig. 4. The SVMs are trained for the unintended DG tripping scenarios seen by the target relay. Simulations have been performed by PSS/E software on a PC with an Intel Xeon W3530C 2.8 GHz CPU. LIBSVM is used to train and test the SVMs [34,43].

5.1. Observation of system dynamic behavior

As mentioned before, the frequency deviation is not a concern in this study and the system frequency has been observed to be robust to the DG tripping events as the frequency deviations have not been significant in an absolute value, typically less than 0.0002pu. This is because the majority of power loss is compensated by the rest of the generators in the system and the strong interconnection with the rest of the US/Canada network, represented by the generator at bus 39.

Fig. 5 shows the apparent impedance trajectory seen by the target relay (R_{25-26}) when a 3-phase fault happens in the middle of the line 26–29, which clears after 0.2 s by tripping the line 26–29 out,

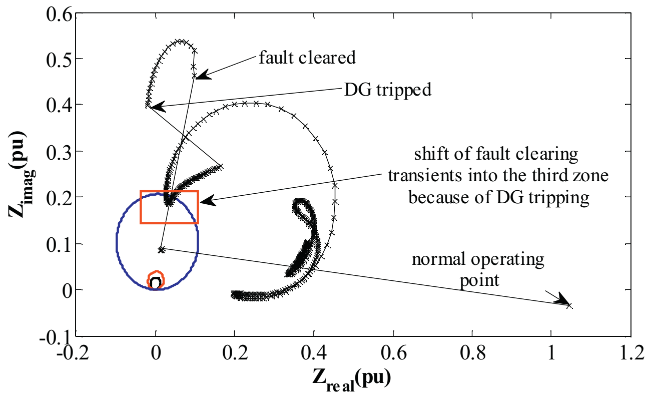

Fig. 5. The apparent impedance trajectory for a DG tripping scenario following a short circuit in the system.

and then 0.27 s later 150 MW PV is tripped. As Fig. 5 shows, the DG tripping has caused the impedance trajectory after the fault clearing to be shifted into the third zone of the relay, which can initiate a false trip as a result. According to measurements from the simulation, the target relay sees the impedance trajectory in its third zone for about 60 cycles which is critically close to issuing a trip signal. Obviously, based on the DG tripped capacity and the instant of DG tripping, the zone interference increases or decreases.

5.2. Creating the training and testing data sets

The SVM training data set consists of different cases (84 cases in total) including: 3 DG tripped capacities (100 MW, 250 MW, and 500 MW), faults on transmission system at different distances in the vicinity of the DG placement, which also includes some points along the lines in the third zone of the target relay, and multiple DG tripping instants following the disturbance. According to the IEEE standard, the anti-islanding schemes should be able to detect all possible islanding conditions and trip DGs within 0.16 to 2 s depending on the level of voltage and frequency variations [8]. The reporting rate of the PMUs is considered 60 phasor per second in a 60 Hz system according to the standard [40]. It should be noted that this is different from the PMU sampling rate on the input signal. The sampling rate might be up to 512 samples per cycle [40]; however, one phasor per cycle is computed and reported by the PMU. Each instance of training includes 2 cycles of data. As mentioned before, the local measurements include V_{bus},$|I_{line}|$,P_{line}, and Q_{line} at the target relay location. Deploying the PMU technology, V_{DG}, P_{DG}, and Q_{DG} at the DG PCC are also available. Note that the voltage phasors measurements include both magnitude and angle. Therefore, considering 1 phasor per cycle reporting rate and length of each instance (2 cycles), the input vector for each instance of training consists of 10 (5×2) features from local measurements and 8 (4×2) from those of WA. Hence, the input vectors for SVM-1 and SVM-2 consist of 10 and 18 features respectively. Considering all

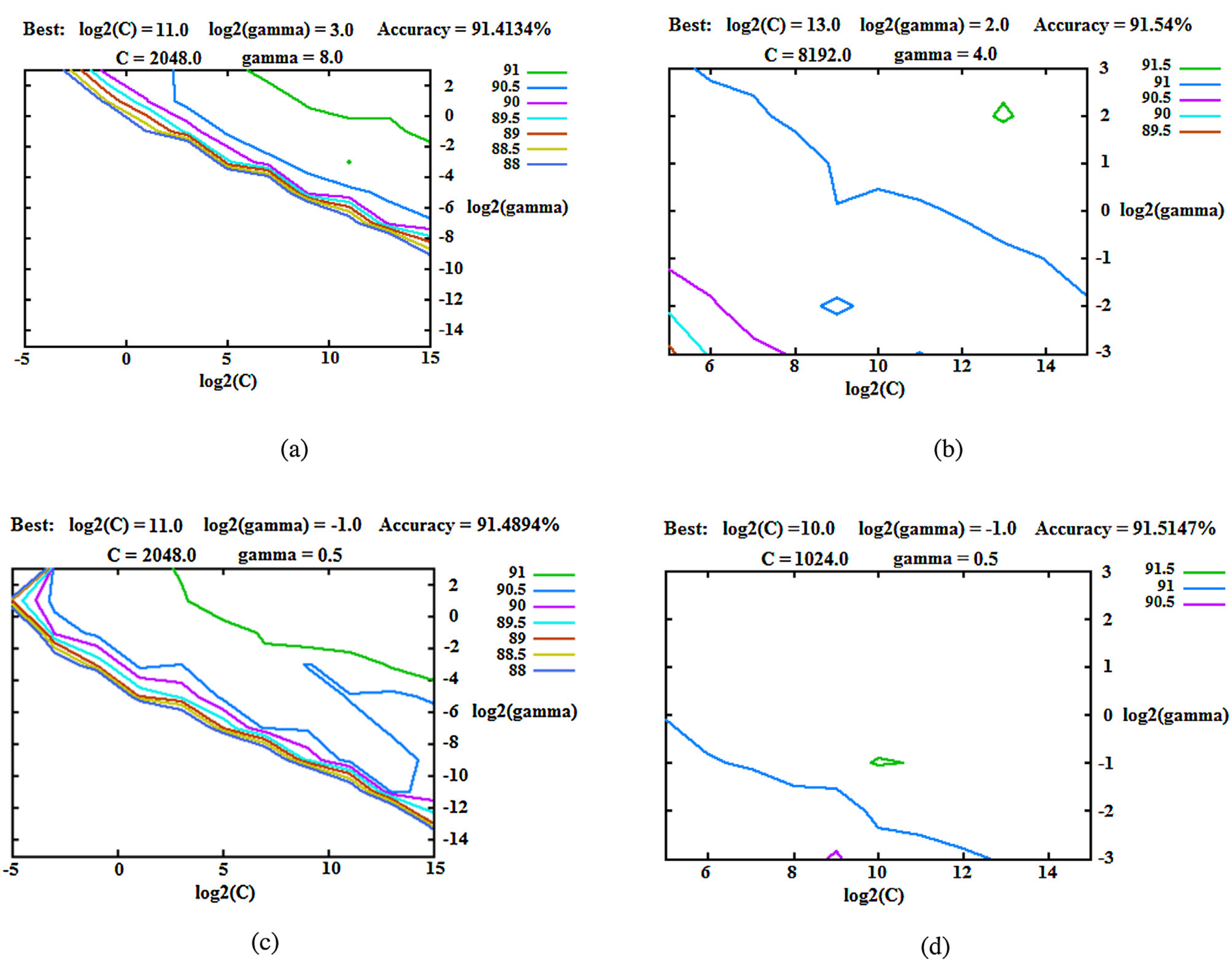

Fig. 6. Interactive grid search using cross-validation for selecting SVMs parameter values; (a)-(b) loos and fine searches on training data-set for SVM-1; (c)-(d) loos and fine search on training data-set for SVM-2.

Table 2
SVMs specifications.

SVM No.	SVM-1	SVM-2
C	8192	1024
γ	4	0.5
No. of Iterations	2112058	98096
No. of SVs	772	753
Testing Accuracy (%)	93.8	97.6
Training Time (s)	39.67	8.76
Testing Time (s)	0.147	0.19

the simulated training cases, i.e. 84 cases of one-and-half seconds system operation time, the training set consists of 3780 instances. The same procedure is taken to create the testing data set. The conditions including DG tripped capacity, fault location and instant of DG tripping, are chosen intentionally different from the training set to assess the performance of the SVMs for unseen scenarios. In total, there are 1692 instances in the testing data set. Different types of DGs and their modeling might cause a change in the measurement values of the selected features and the SVMs should be trained based on the updated values correspondingly.

5.3. SVM parameters selection, training, and testing

The next step is to select SVMs parameters efficiently. It can get time consuming to choose the best parameters in case proper selection cannot be drawn from the available knowledge on the problem and a systematic approach is not taken. An interactive grid search approach has been taken here to evaluate the training generalized accuracy. A five-fold cross validation is implemented by dividing the training data set into 5 subsets of data. It is found that an effective approach to obtain proper values for the pairs of (C, γ) is to try growing their sequence of values exponentially. In order to avoid a complete grid search which is time consuming, first, a bigger incremental step is chosen for the sequence of values, e.g. $C = 2^{-5}$, 2^{-3},. . ., 2^{15} and $\gamma = 2^{-15}$, 2^{-13},. . ., 2^{3}, which is called a loos grid search to find a proper region on the grid of data. When the proper parameter values are found, another search with smaller incremental step is conducted around the values, which is called fine grid search, to find any better value for the parameters. The graphical presentation of generalization contours for the SVMs after a five-fold cross-validation is shown in Fig. 6 in which (a) and(b) and (c) and (d) corresponds to SVM-1 and SVM-2 for loos and fine searches on training grid of data respectively. Although the accuracies for the cases seen in Fig. 6 are not significantly different, it should be noted that this is training accuracy which does not replicate the SVM performance fairly as the class labels are known. To verify the SVM actual performance, it should be evaluated on a data set with unknown labels to observe the testing accuracy. As it could be seen from Fig. 6, proper values for the pairs of (C, γ) for SVM-1 and SVM-2 are determined to be (8192,4) and (1024,0.5) respectively. The parameter selection process has been accomplished in less than 10 min.

Having found the proper parameters values, the SVM-1 and SVM-2 are trained and tested for their corresponding sets of data. Table 2 shows the SVMs specifications and classification accuracy obtained in both cases of using local measurements only (SVM-1), and including WA measurements as well (SVM-2). As it could be seen, the classification accuracy has been increased to a very desirable level when employing WA measurements; however, an acceptable accuracy is still achieved while using local measurements only. This assures the robustness of the method against PMU data unavailability. In addition, the insignificant training and testing time for SVMs as shown in Table 2, infers the easy implementation and practicality of the proposed method.

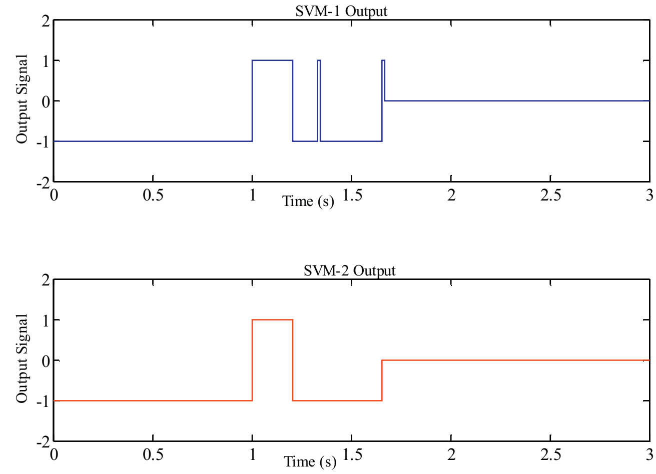

Fig. 7. SVMs output comparison under the DG tripping scenario.

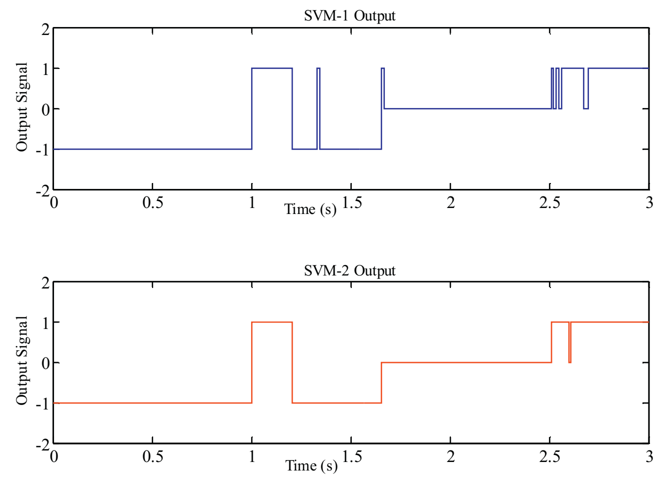

Fig. 8. SVMs output comparison for a fault during the DG tripping scenario.

Fig. 7 compares the outputs of SVM-1 and SVM-2 for a testing DG tripping scenario. The testing scenario is that a 3-phase fault happens on x = 0.3 of the line 26–29 at t = 1 s, and cleared at t = 1.2 s by tripping this line out. As a consequence of miss-detection of PV interconnection relays on the distribution side, at t = 1.65 s, 250 MW PV is tripped unintentionally. The signal values of "1", "0", and "−1" represent the fault, DG tripping, and other status respectively. As it could be seen both SVMs have classified the instances well. The temporary spikes seen in SVM-1 output (at t = 1.35 s or t = 1.65 s) are as a result of miss-classification; however, they do not affect the relay operation since they are not persistent.

As mentioned before, employing the PMU measurements from PCC improves the accuracy of the SVM especially under more complicated scenarios of classification. To verify this, the above mentioned scenario has been complicated by the occurrence of a second fault on x = 0.4 of the line 26–28 at t = 2.5 s when the system is already experiencing the stress caused by previous fault and subsequent DG tripping. The SVMs' performance under this scenario is illustrated in Fig. 8. As it could be seen, SVM-2 has classified the instances with a better accuracy compared to SVM-1, i.e. lower number of misclassifications, especially during the second fault detection. Fig. 9 compares the proposed method output, i.e. trip/block signal in Fig. 2, with the conventional relay pickup on the above mentioned testing scenario. As it could be seen, following DG tripping, the distance element of the target relay backup zone (zone 3) has picked up from t = 1.8 s to the end while the proposed

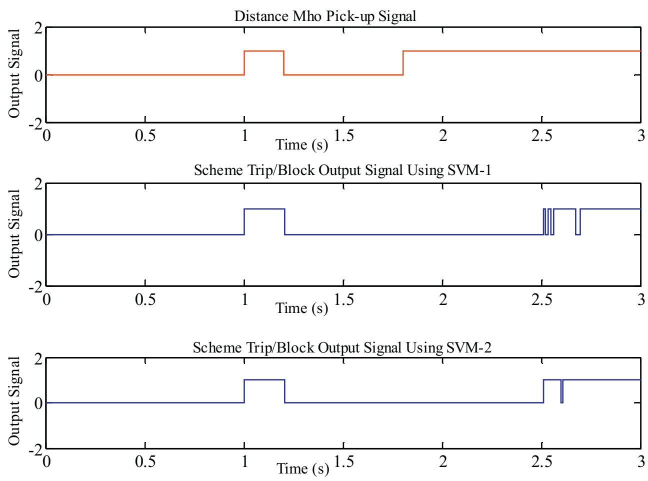

Fig. 9. Comparison of the proposed method output with the conventional distance pickup.

Table 3
SVMs performance comparison.

Type of Instance		Fault	DG Tripping	Others
Actual No. of Instances		263	897	532
No. of Detected Instances	SVM-1	284	913	495
	SVM-2	270	891	531
No. of Correctly Detected Instances	SVM-1	240	860	488
	SVM-2	251	874	527
No. of Incorrectly Detected Instances	SVM-1	44	53	7
	SVM-2	19	17	4
Accuracy (%)	SVM-1	91.3	95.9	91.7
	SVM-2	95.4	97.4	99.1

method blocks the relay operation during DG tripping interference and unblocks it at t = 2.5 s when the second fault happens.

Table 3 summarizes and compares the performance of the SVMs in classifying the instances. As it could be seen from Table 3, the numbers of correctly detected cases of faults, DG tripping, and others are higher when deploying SVM-2. In other words, the protection dependability and security has been better maintained when using WA measurements. The accuracy in Table 3 is the ratio of the number of correctly detected instances of a type (e.g. fault) to the actual number of instances of that type. As mentioned before, the fault instances include complicated fault scenarios, i.e. a second fault happening on transmission side following DG tripping event, to test the dependability of the proposed method in addition to the security. To have an estimate on the proposed method's dependability for normal fault situations on the transmission side, i.e., faults which are not following the DG tripping event, a total 117 of such fault cases were run and SVM-1 and SVM-2 performed detection with 99.1% and 100% accuracy respectively.

6. Conclusion

The main contributions of this paper are as follows:

- A SVM-based protection scheme which prevents maloperation of distance relays in unintended DG tripping scenarios is proposed.
- WA measurements have been used in addition to local measurements to increase the SVM's classification accuracy and that of the protection scheme consequently. The proposed scheme is robust against PMU data loss or unavailability.
- Unlike conventional blocking schemes, the proposed protection scheme not only blocks the relay following the interference of a DG tripping scenario with distance coordination but also detects a fault if it happens during the blocking period and unblocks the relay to operate properly.
- Since the proposed scheme is easily and quickly trainable, it is applicable to various possible practical system operation scenarios and gives significant selectivity.

In summary, deploying the WA measurements infrastructure not only improves the scheme accuracy but also makes it independent of the aggregated DG location in the system. The proposed scheme could be implemented in combination with other protection schemes such as power swing blocking to help maintaining power system protection dependability and security.

As future steps of this work, we will study the probable DG tripping on multiple PCCs in a network and how to upgrade the proposed method for such a case by proper SVM training and WA measurement deployment.

References

[1] M. Tasdighi, Residential microgrid scheduling based on smart meters data and temperature dependent thermal load modeling, IEEE Trans. Smart Grid 5 (January (1)) (2014) 349–357.
[2] IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems, IEEE Standard (2003)1547–2003.
[3] IEEE Recommended Practice for Utility Interface of Photovoltaic (PV) Systems, IEEE Standard 929–2000, Jan. 2000.
[4] Standard for Inverters, Converters, and Controllers for Use in Independent Power Systems, UL Standard 1741, Underwriters Laboratories Inc., May 1999.
[5] New York State Standardized Interconnection Requirements, Application Process, Contract and Application Forms for New Distributed Generators, 300 Kilovolt-Amperes or Less, Connected in Parallel with Radial Distribution Lines, New York State Public Service Commission, Nov. 2000.
[6] Rule 21-Generating Facility Interconnections, California Public Utilities Commission, 2000.
[7] Texas Public Utility Commission Requirements for Pre-Certification of Distributed Generation Equipment by a Nationally Recognized Testing Laboratory, Public Utility Commission of Texas, Feb. 2001.
[8] B. Matic-Cuka, M. Kezunovic, Islanding detection for inverter-based distributed generation using support vector machine method, IEEE Trans. Smart Grid 5 (November (6)) (2014) 2676–2686.
[9] S.P. Chowdhury, et al., Islanding protection of active distribution networks with renewable distributed generators: a comprehensive survey, Electr. Power Syst. Res. 79 (June (6)) (2009) 984–992.
[10] J. Peças Lopes, Integration of dispersed generation on distribution networks—impact studies, in: Proceedings of the IEEE Winter Meeting, N.Y, February, 2002.
[11] C. Zheng, M. Kezunovic, Distribution system voltage stability analysis with wind farms integration, in: IEEE PES 42nd North American Power Symposium (NAPS), Arlington, USA, September, 2010.
[12] F. Katiraei, J.R. Aguero, Solar PV integration challenges, IEEE Power Energy Mag. 9 (May–June (3)) (2011) 62–71.
[13] M. Henderson, Impacts of transmission system contingencies on distributed generation -overview, ISO New-England DG Forecast Group Meeting (2013) http://www.cancerresearchuk.org/aboutcancer/statistics/cancerstatsreport/.
[14] R. Walling, DER Impacts on Bulk Grid Dynamic Performance, EPRI-PJM Inverter Interconnections Workshop, April, 2012. http://mydocs.epri.com/docs/publicmeetingmaterials/4-17-2012/05B-DER-Impacts-on-Bulk-Grid-Dynamic-Performance-Reigh-Walling.pdf.
[15] M.H. Bollen, F. Hassan, Integration of Distributed Generation in the Power System, first-ed., Wiley-IEEE Press, 2011.
[16] J.A. Peças Lopes, et al., Integrating distributed generation into electric power systems: a review of drivers, challenges and opportunities, Elect. Power Syst. Res. 77 (9) (2007) 1189–1203.
[17] R.A. Walling, N.W. Miller, Distributed generation islanding-implications on power system dynamic performance, IEEE Power Eng. Soc. Summer Meet. 1 (2002) 90–96, 25–25 July.
[18] NERC Standard for Generator Frequency and Voltage Protective Relay Settings, NERC Standard PRC-024-2.
[19] S. Eftekharnejad, et al., Impact of increased penetration of photovoltaic generation on power systems, IEEE Trans. Power Syst. 28 (May (2)) (2013) 893–901.
[20] Adjustment Period and Real-Time Operation, ERCOT Nodal Protocol, http://www.ercot.com/content/wcm/libraries/105131/September_1__2016_Nodal_Protocols.pdf.
[21] NERC Standard for Transmission System Planning Performance Requirements, NERC Standard TPL-001-4.
[22] D.M. MacGregor, et al., Automatic relay setting, J. Electr. Electron. Eng. 21 (3) (2002) 169–179.

[23] A. Gopalakrishnan, et al., Simulating the smart electric power grid of the 21st Century—bridging the gap between protection and planning, Western Protective Relay Conference (2013).
[24] M. Kezunovic, C. Zheng, Monitoring power system dynamic performance using synchronized sampling, in: CIGRE International Conference on Advances in Power System Control, Operation and Management (APSCOM), St. Petersburg, Russia, April, 2008, 2008.
[25] A. Ukil, Intelligent Systems and Signal Processing in Power Engineering, Springer, Heidelberg, Berlin, Germany, 2007, pp. 161–226.
[26] B. Ravikumar, et al., An approach using support vector machines for distance relay coordination in transmission system, IEEE Trans. Power Deliv. 24 (January (1)) (2009) 79–88.
[27] B. Ravikumar, et al., Comparison of multiclass SVM classification methods to use in a supportive system for distance relay coordination, IEEE Trans. Power Deliv. 25 (July (3)) (2010) 1296–1305.
[28] P.K. Dash, et al., Fault classification and section identification of an advanced series-compensated transmission line using support vector machine, IEEE Trans. Power Deliv. 22 (January (1)) (2007) 67–73.
[29] U.B. Parikh, et al., Combined wavelet-SVM technique for fault zone detection in a series compensated transmission line, IEEE Trans. Power Deliv. 23 (October (4)) (2008) 1789–1794.
[30] K. Seethalekshmi, et al., A classification approach using support vector machines to prevent distance relay maloperation under power swing and voltage instability, IEEE Trans. Power Deliv. 27 (July (3)) (2012) 1124–1133.
[31] M. Tasdighi, M. Kezunovic, Automated review of distance relay settings adequacy after the network topology changes, IEEE Trans. Power Deliv. PP (99) (2016), http://dx.doi.org/10.1109/TPWRD.2016.2524654, 1–1.
[32] M. Tasdighi, M. Kezunovic, Impact analysis of network topology change on transmission distance relay settings, in: IEEE Power & Energy Society General Meeting, Denver, CO, 26–30, July 2015, 2015, pp. 1–5.
[33] V. Vapnik, The Nature of Statistical Learning Theory (Information Science and Statistics), Springer, New York, NY, USA, 1999, Dec.
[34] J. Weston, C. Watkins, Multi-class support vector machines, Dept. Comput. Sci., Royal Holloway, Univ. London, Egham, U.K., Tech. Rep. CSD-TR-98-04 (1998). [Online]. Available: http://citeseer. ist.psu.edu/article/weston98multiclass.html.
[35] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines., ACM Transactions on Intelligent Systems and Technology, 2, 27, 1–27:27 (2011). [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm.
[36] S.S. Keerthi, C.-J. Lin, Asymptotic behaviors of support vector machines with Gaussian kernel, Neural Comput. 15 (7) (2003) 1667–1689.
[37] H.-T. Lin, C.-J. Lin, A Study on Sigmoid Kernels for SVM and the Training of Non-PSD Kernels by SMO-Type Methods, Technical report, Department of Computer Science, National Taiwan University, 2003 [Online]. Available http://www.csie.ntu.edu.tw/~cjlin/papers/tanh.pdf.
[38] J.G. Slootweg, et al., Modeling new generation and storage technologies in power system dynamics simulations, in: Proceedings IEEE Summer Meeting, Chicago, July, 2002, 2002.
[39] M.M. Eissa, et al., A novel back up wide area protection technique for power transmission grids using phasor measurement unit, IEEE Trans. Power Delivery 25 (January (1)) (2010) 270–278.
[40] IEEE Standard for Synchrophasor Data Transfer for Power Systems, in IEEE Std C37.118. 2–2011 (Revision of IEEE Std C37.118-2005).
[41] R. Hasan, et al., Analyzing NASPInet data flows, Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES, Seattle, WA, 2009, pp. 1–6.
[42] Power System Test Case Archive, Univ. Washington, Dept. Elect. Eng., 2013. [Online]. Available: https://www.ee.washington.edu/re-search/pstca/index.html.
[43] C.-W. Hsu, C.-C. Chang, C.-J. Lin. A practical guide to support vector classification. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf.