21, rue d'Artois, F-75008 PARIS
http : //www.cigre.org

CIGRE US National Committee
2016 Grid of the Future Symposium

Predicting Power Grid Component Outage In Response to Extreme Events

R. ESKANDARPOUR, A. KHODAEI
University of Denver
USA

S. BAHRAMIRAD
ComEd
USA

M. BOLLEN
Luleå Univ of Technology
Sweden

SUMMARY

An accurate forecast of the power system component outages during extreme events is an essential task to improve pre-event system preparedness and post-event system recovery and accordingly minimize the undesired aftermath of these events. A machine learning method, based on Support Vector Machines (SVM), is proposed in this paper as a viable approach to forecast the components which can potentially fail during an anticipated extreme event. In particular, a cubic kernel SVM is proposed to classify between operational and damaged components after the extreme event based on the event characteristics. The extreme event can be of the nature of a weather event or a natural disaster, where in either case the proposed approach is capable of developing suitable prediction models. The proposed method can be trained on historical data of the past extreme events. The performance of the proposed method in effectively predicting potential component outages is validated using two defined metrics, namely precision and recall. Numerical studies indicate that the proposed method can be used to effectively predict the outages.

KEYWORDS

Power system resilience, extreme event, machine learning.

Rozhin.Eskandarpour@du.edu

1. INTRODUCTION

Extreme events, such as hurricanes, snowstorms, floods, earthquakes, etc., cause outages in critical lifeline systems and result in inconvenience for residents living in disaster areas [1]. The electricity infrastructure, as a critical lifeline system, which is spread over wide geographical areas to transfer bulk power from centralized power plants to distributed load centres, is not an exception to this impact. Considering the large impacts of extreme events on electric power systems, including local and national losses in range of billions of dollars every year [2], a viable prediction, response, and recovery is of significant importance. In other words, the power system needs to become more resilient in response to these events. The concept of resilience in power systems, initially introduced in [3], determines the resistance of the power grid and its ability to withstand extreme changes. Power system resilience has turned into a progressively essential affair as the frequency and the intensity of extreme whether events have significantly grown in recent years. The considerable consequences of extreme events on the electricity infrastructure, which is spread over a broad geographical area and hence vulnerable to be largely impacted, shows the value and importance of an efficient forecasting of the potential damages to power system components which would accordingly enable efficient response and recovery schemes.

Traditionally, the power system component outages were predicted holistically (i.e. at the system level but not at the component level) or were estimated using probabilistic approaches following a predefined probability distribution function [4, 5]. These methods suffer from several drawbacks, in which the holistic methods are not useful in managing the grid components and the probabilistic methods may not be accurate and could further vary for different regions and events. Machine learning approaches, on the other hand, have shown a great performance on learning from and making predictions on existing data. These approaches build a model from an example training set of observations without explicitly defining the probabilistic model, and predict data-driven decisions as outputs. One of the challenges in machine learning approaches is to have adequate number of samples for training to extract necessary features to train the model. As for extreme events, these data can be easily derived from past events.

In this paper, a machine learning method is proposed to determine the power system component outages in response to an anticipated extreme event. The rest of the paper is organized as follows: Section 2 presents the problem statement and proposes the machine learning method for outage prediction; Section 3 presents simulation results on a test system; and Section 4 concludes the paper.

2. MACHINE LEARNING METHOD FOR OUTAGE PREDICTION

The state of each component in the power system in the path of an upcoming hurricane can be considered as (a) damaged, which means the component is on outage, or (b) operational, which means the component is in service. The path and the intensity of the hurricane can be anticipated from weather agencies. In order to classify the damage state of the power system components, different features can be extracted from historical data. In this paper, we explore two main features of the wind speed and the distance of the each component from the center of the hurricane. A Support Vector Machine (SVM) [6] is used for this purpose and to further determine the decision boundary between the damaged and operational data points.

Given a set of training examples, an SVM classifies them into two classes by finding the best hyperplane that separates training examples of one class from the other class. The best hyperplane is defined as the hyperplane with a clear gap that is as wide as possible. Figure 1 shows the support vectors and optimal hyperplane in a separable two class classification of SVM.

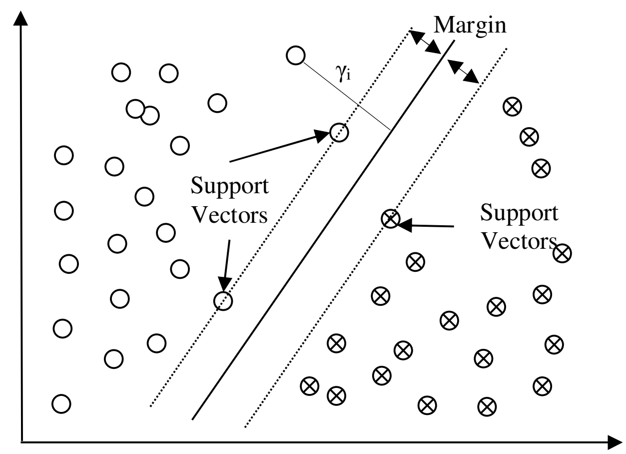

Figure 1. Support vectors and optimal margin in SVM

The data for training is a set of points x_i ($x_i \in R_D$) along with their categories y_i ($y_i = \pm 1$), the classification task can be written as:

$$h_{w,b}(x) = \text{sgn}\left(w^T x + b\right). \tag{1}$$

where w is the normal vector to the hyperplane separating training examples, $|b|/||w||$ is the perpendicular distance from the hyperplane to the origin, and *sgn* is the sign function, i.e., sgn(z) = 1 if $z \geq 0$, and sgn(z) = −1 otherwise. $h(x)$ is the output of the classifier with the aim of $h(x_i)$=1 if y_i=+1 and $h(x_i)$= −1 otherwise. We can then define a large functional margin that representing a confident prediction as:

$$\hat{\gamma}_i = y_i\left(w^T x_i + b\right) \tag{2}$$

We can define geometric margin (shown in Figure 1) as:

$$\gamma_i = \frac{\hat{\gamma}_i}{\|w\|} \tag{3}$$

Given a training set S = {(x_i, y_i); i=1,…,m}, the geometry margin of the decision hyperplane (w,b) with respect to S is defined to be the smallest functional margins of individual training examples, as:

$$\gamma = \min_{i=1,\dots,m} \gamma_i \tag{4}$$

Sine we are maximizing the functional margin of the decision hyperplane (w,b), we can maximize the geometry margin (as it is scale invariant), while minimizing $||w||^2$. Then, the optimal hyperplane parameters (w,b) can be found by optimization problem:

$$\begin{aligned} &\min_{\gamma,\omega,b} \frac{1}{2}\|w\|^2 \\ &s.t. \quad y^{(i)}\left(w^t x^{(i)} + b\right) \geq 1, \qquad i = 1,......, m \end{aligned} \tag{5}$$

This is a quadratic programming problem, which can be solved by a Lagrange duality. Solving the duality, the final hyperplane only depends on the support vectors (i.e., samples points that are in the margin) and SVM needs to find only the inner products between the test samples and the support vectors (of which there is often only a small number).

In case that the training data cannot be separated by a hyperplane (which commonly happen, especially in case of the hurricane data), SVM can use a soft margin. This can be solved by a penalty parameter c and a regularization (often L1 or L2) as follows:

$$\begin{aligned} &\min_{\gamma,\omega,b} \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{m}\varepsilon_i \\ &s.t. \quad y^{(i)}\left(w^t x^{(i)} + b\right) \geq 1, -\varepsilon_i, \qquad i = 1,......, m \\ &\qquad \varepsilon_i \geq 0, \qquad i = 1,......, m \end{aligned} \tag{6}$$

In other words, training examples can have a margin less than one, and if an example has functional margin $1\text{-}\varepsilon_i$ (with $\varepsilon_i > 0$), the objective function is increased by $c\varepsilon_i$. Finding a proper value of c depends on the shape of classes, which are often unknown. Therefore, c is often found by testing the performance of the classifier on a validation set.

The idea of maximum-margin hyperplane, which is discussed above, is based on the assumption that training data are linearly separable, which is not the case in many practical applications. In order to apply SVM to nonlinear data, kernel methods [6] can be used. The idea of kernel method (kernel trick) is to map input features to higher demotions that can be linearly separable and fit the maximum-margin hyperplane in the transformed feature space. Kernel trick simply states that for all x_1 and x_2 in the input space X, a certain function $k(x_1,x_2)$ can be replaced as inner product of x_1 and x_2 in another space. For example a polynomial kernel can be defined as:

$$k\left(x_i, x_j\right) = \left(x_i . x_j\right)^d \tag{7}$$

Training samples may still be non-linearly separable in the transformed feature space. Therefore, multiple SVM are trained with various kinds of kernels (e.g. polynomial with different degrees, Gaussian, etc.) and the best kernel is found imperially from the result on a validation set. The role of the penalty parameter can be also important in finding the best setting for the problem.

To evaluate the performance of the classifier, usually a subset of historical data is reserved as the validation/test set. Reporting the general accuracy of prediction cannot be sufficient as number of samples may not balance in the test set. The F_1-Score is a common and reliable measure of classification performance [7] which will be tested on the test historical data:

$$F_1 = \frac{2PR}{(P+R)} \tag{8}$$

where P and R represent precision and recall metrics, respectively. Precision is defined as the number of correctly predicted outages divided by the total number of *predicted* outages, and the recall is defined as the number of correctly predicted outages divided by the total number of *actual* outages. Precision can be seen as a measure of a classifier exactness and recall can be thought of as a classifiers completeness. A higher value of the F_1-Score, which is a number between 0 and 1, indicates a better forecasting and justifies the viable performance of the existing decision boundary.

3. CASE STUDY

As historical data for the past extreme events at component level are limited, we have generated 300 samples of each component state following a normal distribution function with a small Gaussian noise so that the data can be distinguishable. The samples belong to two classes of components with high probability of failure and components that can survive the extreme event. The features are normalized to [0, 1] based on the maximum considered values of wind speed and distance. These samples are shown in Figure 2.

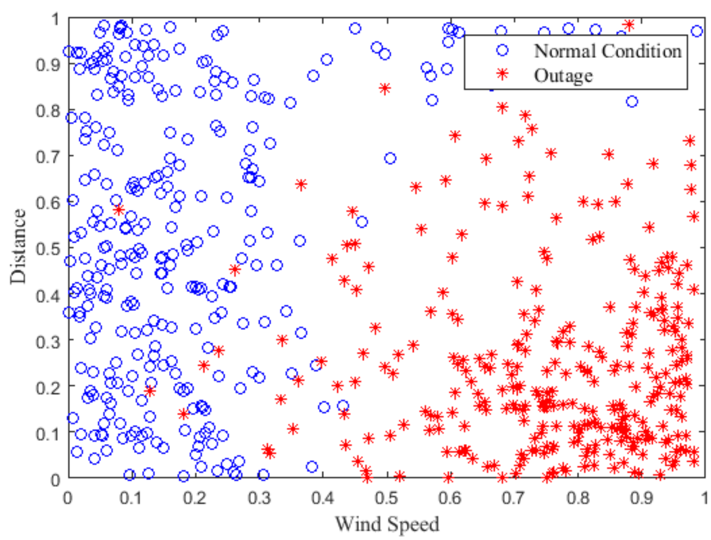

Figure 2. Generated samples for two different classes

A k-fold cross validation (k=5) is performed to measure the performance of the proposed method, where the generated data is split into training and validation subsets. During the training of the system, the SVM model is only trained on the training subset and validated on the subset that is not trained on. Different kernels (linear, polynomial Quadratic and Cubic) with different range of penalty parameter (c=0.01, 0.1, 1, 10, 100) are trained. Among the trained SVM, polynomial Cubic kernel with c=1 had the best overall classification accuracy on the 5 fold of validation set. The final result is the average accuracy over all k folds. The average overall classification accuracy of the proposed classification model is 96.0% with the average F_1-Score of 95.97% for predicting outage. Table 1 shows the confusion matrix of classifying components into two classes of outage (having high probability of failure) and normal condition based on the distance to the center of extreme event and the wind speed that they can withstand. As observed, the proposed method can classify the outage components from normal condition with high accuracy. The proposed model is a general framework that can be improved by extracting more features (i.e. different types of components, etc.) and can be easily adopted to historical data if the component-level outage data are available.

Table 1. Confusion Matrix of classifying system component during extreme event

		Predicted	
		Normal	**Outage**
Actual	**Normal**	96.7%	3.3%
	Outage	4.7 %	95.3%

4. CONCLUSION

Predicting power system outages at the component level is an important factor in scheduling power system response and recovery against extreme events. In this paper, a machine learning based outage prediction model was proposed to determine the probable outage of power system components based on historical event data and specific event characteristics. A case study on synthetically generated data showed that the proposed model can effectively predict outages while offering a great generalization capacity for new samples in the test subset. The generated data was aimed to study the effect of hurricanes on the system, but the proposed model is applicable to a variety of extreme events, and also able to consider a wide range of other features in addition to hurricane speed and component distance.

BIBLIOGRAPHY

[1] H. E. Moore, F. L. Bates, M. V. Layman, and V. J. Parenton, “Before the wind: A study of the response to Hurricane Carla” (National Academy of Science, Washington, DC, USA, Tech. Rep., 1963).

[2] Executive Office of the President, “Economic Benefits of Increasing Electric Grid Resilience to Weather Outages” (August 2013).

[3] C. S. Holling, “Resilience and stability of ecological systems” (Annual Review of Ecology, Evolution, and Systematics, pages 1–23, 1973).

[4] A. Arab, A. Khodaei, S. K. Khator, K. Ding, V. A. Emesih, and Z. Han, “Stochastic pre-hurricane restoration planning for electric power systems infrastructure” (IEEE Transactions on Smart Grid, volume 6, no. 2, pages 1046–1054, 2015).

[5] A. Arab, A. Khodaei, Z. Han, and S. K. Khator, “Proactive recovery of electric power assets for resiliency enhancement” (IEEE Access, volume 3, pages 99–109, 2015).

[6] C. Cortes, and V. Vapnik, “Support-vector networks” (Machine learning, volume 20, no. 3, pages 273-297, 1995).

[7] R. O. Duda, and P. E. Hart. “Pattern classification and scene analysis” (volume 3, New York: Wiley, 1973).