Support-Vector-Machine-Based Proactive Cascade Prediction in Smart Grid Using Probabilistic Framework

Sudha Gupta, *Member, IEEE*, Ruta Kambli, Sushama Wagh, *Member, IEEE*, and Faruk Kazi, *Member, IEEE*

***Abstract*—The worldwide major blackout events of power network are highlighting the need for technology upgradation in traditional grid. One of the major upgradations required is in the area of early warning generation in case of any grid disturbances such as line contingency leading to cascade failure. This paper proposes a proactive blackout prediction model for a smart grid early warning system. The proposed model evaluates system performance probabilistically, in steady state and under dynamical (line contingency) state, and prepares a historical database for normal and cascade failure states. A support vector machine (SVM) has been trained with this historical database and is used to predict blackout events in advance. The key contribution of this paper is to capture the essence of the cascading failure using probabilistic framework and integration of SVM machine learning tool to build a prediction rule, which would be able to predict the scenarios of the blackout as early as possible. The proposed model is validated using the IEEE 30-bus test-bed system. Proactive prediction of cascade failure using the proposed model may help in realizing the grid resilience feature of smart grid.**

***Index Terms*—Cascade failure, cumulative distribution function (CDF), Gaussian distribution, probability density function (PDF), support vector machine (SVM).**

I. Introduction

CASCADE failure is a mechanism where failures propagate to cause large blackouts of the electrical power system. These blackouts impact on the social, economic, and industrial growth of the country. Some of the major worldwide blackouts recorded and reported in the literature [1] are the following: July 2 and August 10, 1996, blackout in the U.S.; August 14, 2003, blackout in North America; September 28, 2003, blackout in Italy; November 4, 2006, blackout in Europe; and the two consecutive severe blackouts on July 30 and 31, 2012, in the Indian power grid [2], [3].

The severity of such incidences has motivated researchers to understand and analyze past blackout events and propose a real-time proactive prediction model to prevent massive blackouts. The real-time exchange of data and information needed for early warning system is made possible by the recent advancements in communication technology and sensors [4]–[10] in smart grid.

To understand cascading effect in the traditional grid, a probabilistic load flow model was proposed and compared with a deterministic model [11], [12] in terms of density function. Impact of network uncertainties in the power system was modeled in [13] and [14] using distribution factor concept as a function of linear power injections. The statistical process monitoring is a powerful tool for detecting faults in industrial systems [15]. A stochastic Markov model was used in [16] to capture the progression of cascading failure and its time span. Smart grid was modeled as interdependent complex networks in [17], and cascade failure was analyzed using percolation theory. The increased blackout risk analysis was carried out by [18] using autocorrelation techniques. The analysis results showed that correlation significantly increased before reaching critical line. A probabilistic-regeneration-based approach was proposed in [19] to model the dynamics of cascading failures in power grids using Monte Carlo (MC) simulations. A probabilistic framework for evaluation of smart grid resilience of cascade failure was explored in [20], and the simulation result showed that statistical analysis of probabilistic power flows model is useful for evaluation of cascade failure.

The next-generation power grid, i.e., the smart grid [21], demands high reliability and robustness against cascading failure. The prediction based on stochastic methods with MC, Markov model, and percolation theory is computationally exhaustive and not explored much for proactive prediction. Integration of computational intelligence (CI) with traditional blackout prediction techniques can provide a feasible solution to analyze current, past, and future performance of the grid and increase the reliability of the power system. Researchers are using CI methods to solve many challenging problems [21], [22] of the traditional grid and contributing to a complete realization of smart grid. These intelligent technologies build offline patterns for each operating condition of the grid and monitor an online pattern of the grid status. A classifier compares the previously learned patterns with the pattern generated online in order to

Manuscript received March 4, 2014; revised June 22, 2014; accepted August 6, 2014. Date of publication October 3, 2014; date of current version March 6, 2015. This work was supported in part by the Center of Excellence in Complex and Nonlinear Dynamical Systems, Veermata Jijabai Technological Institute, Mumbai, India, under TEQIP-II subcomponent (1.2.1).

S. Gupta is with the Electrical Engineering Department, Veermata Jijabai Technological Institute, Mumbai 400 019, India, and also with K. J. Somaiya College of Engineering, Mumbai 400 077, India (e-mail: sudhagupta@somaiya.edu).

R. Kambli, S. Wagh, and F. Kazi are with the Electrical Engineering Department, Veermata Jijabai Technological Institute, Mumbai 400 019, India (e-mail: kambliruta@gmail.com; srwagh@vjti.org.in; fskazi@vjti.org.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2014.2361493

classify the grid operating condition and identify the status of the grid.

Most of the researchers used machine learning tools in the power grid to solve fault identification and localization purposes. In [23]–[25], the artificial neural network (ANN) or pattern recognition techniques were employed for faults classification in induction motors (IMs). An artificial intelligence technique is used as an unsupervised classification technique in [26] for the detection and diagnosis of faults in IM. An energy management system for the optimal operation of smart grids and microgrids was proposed in [10] using fuzzy clustering neural networks (NNs) combined with optimal power flow. They used genetic algorithms and fuzzy clustering as a training algorithm.

Machine learning techniques for transforming historical electrical grid data into models to predict the risk of failures for components and systems were explored in [27]. ANN techniques are based on observed error minimization principle, which gives local optimal solution, low convergence rate, and low generalization with less number of samples [28]. Compared with ANN, support vector machine (SVM) is formulated as a quadratic programming problem and gives global optimal solution [29]. The solutions provided by SVMs are theoretically elegant, computationally efficient, and very effective to handle many practical problems.

An SVM machine learning tool was used in [28] for the identification of fault type and location in power distribution system with distributed generation. SVM-based smart relays for mitigation of future blackouts were proposed by [30] to detect location of the fault in the power system.

This paper is a contrast with the earlier blackout work where data and estimated model parameters were derived from MC simulations, and the MC simulated failures were predicted. Database has been created for this paper from the proposed probabilistic model, simulated and verified on IEEE 30-bus test bench system, and used as a historical database. The SVM model has been trained on the basis of this database for blackout prediction. ANN and SVM machine learning tools were used by most of the researchers for identification and location of faults in the power system. This paper proposes a new model where SVM tool is used for proactive blackout prediction in the power system. The methodology used in this paper is capable of analyzing transmission line contingency with the help of probabilistic framework, and at the same time, it can also compare online statistical data with the stored blackouts historical data and predict possible blackout.

This paper is organized as follows. Section II presents power flow model under normal and perturbed conditions of the grid. Theory of SVM is described in Section III. The proposed model is explained in Section IV in two phases. Phase A is an offline probabilistic framework for the generation of historical database (normal and blackout). Phase B explains the modeling of SVM and online state classification. Section V presents a case study to validate the proposed proactive cascade prediction model using IEEE 30-bus test-bed system with modeling and simulation. Some open research issues are discussed in Section VI along with conclusions.

II. Grid Power Flow Analysis

Cascade prediction in a power grid requires study and analysis of the power flow, both in normal operating conditions and under perturbed conditions.

A. Power Flow Under Normal Grid Working

A power grid is a complex network where generator and load buses can be assumed as nodes and transmission lines transformers as links. The net power injected into a node is equal to the total amount of power flowing to neighboring nodes through transmission lines or transformers. The angles δ_i and δ_k are the voltage phase angles at node i (sending) and k (receiving) link, respectively; and X_{ik} is the series reactance of the link between nodes i and k. In general, the active power flow over a line connected to nodes i and k [31] can be defined as

$$P_{ik} = -P_{ki} = \frac{|V_i||V_k|}{X_{ik}} \sin(\delta_i - \delta_k). \quad (1)$$

For DC power flow, the voltage magnitude at all nodes is maintained at 1 per unit. Furthermore, as the system synchronization is always maintained under normal operating conditions, the angular difference between two neighboring nodes is very small. Hence, $(\delta_i - \delta_k)$ being very small, $\sin(\delta_i - \delta_k)$ is approximately equal to $(\delta_i - \delta_k)$, which will modify (1) to

$$P_{ik} = \frac{(\delta_i - \delta_k)}{X_{ik}}. \quad (2)$$

Equation (2) determines power flow under a specific operating condition with initial installation capacity (IC), and the accuracy of the model depends on the accuracy of the input data. Any changes in input data will change power flow analysis. As soon as load flow exceeds its IC, the line will trip. Tripping of one line will change corresponding node phase angles δ_i and δ_k; hence, it redistributes the power flow in the grid. Power redistribution may further change grid variables such as current, voltage, active/reactive power, and phase angle and initiate cascading failure. For the stability of the grid, prior knowledge of the line outage is required. The multivariate phenomena of deterministic load flow increase complexity to predict the line outage in advance.

The probabilistic load flow analysis of normal grid operation (base case, with load flow under IC) considers the normal (symmetrical/Gaussian) distribution of power flow. Under the cascading failure, probabilistic load flow takes uncertainties into account, such as probability of a line flow being greater than the capacity of the transmission line, and the grid variables are considered as random variables. In such a scenario, the probability distribution curves will no longer be symmetrical. Hence, the behavior of the grid (switching from normal to cascading failure) can be analyzed in depth with the probability distribution curves and statistical parameters such as probability density function (PDF), cumulative distribution function (CDF), mean, variance, and higher order moments.

A probabilistic model [32] can be designed on the basis of a time instance model for cascade prediction and a time period model for power flow analysis. The time period model over a certain period T is used to calculate and analyze load duration curve, mean, variance, and CDF function for (x, T), where x is the input data (percentage power loading in the transmission lines). The time period model over a certain period T of a transmission line can be described by Gaussian distribution [33] under normal grid operation. Thus,

$$\text{PDF} = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3}$$

$$\text{CDF} = F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{4}$$

where

σ = standard deviation of random variable;

σ^2 = variance of random variable;

x = random variable (current, voltage, power, phase angle);

μ = mean value of random variable;

$f(x)$ = probability density function (PDF).

In (3), the line with highest PDF value in the observed time scale (sample) is used in the probabilistic framework for prediction of the next line outage, and for each sample time T, its mean and variance are calculated as

$$\mu = \frac{1}{N} \sum_{n=1}^{N} x_n \tag{5}$$

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^2 \tag{6}$$

where $N = (1, 2, \ldots, n)$ represents the number of samples. Equations (3)–(6) are used in this paper to capture the statistical information of the grid parameters under normal working state for a time period T and used as normal grid feature vectors.

B. Power Flow Under Perturbed Condition of Grid

Analysis of power flow on the transmission line in perturbed condition such as line tripping provides the information of random variables. Critical transmission line analysis is performed in [34] to find which lines may have negative maximum impact on the grid when they are removed from the system. Tripping of the line will change grid variables and symmetry of probability distribution curves. Hence, to know the lack of symmetry in probability distribution curve, higher order moments [33] such as skewness (third moment) and kurtosis (forth moment) have been used in this paper, i.e.,

$$\text{Skewness} = \frac{\sum_{i=1}^{N} (x_i - \mu)^3}{(N-1)s^3} \tag{7}$$

$$\text{Kurtosis} = \frac{\sum_{i=1}^{N} (x_i - \mu)^4}{(N-1)s^4} \tag{8}$$

where s is a standard deviation. The probabilistic model predicts line outage based upon current line status (line having maximum PDF), and when CDF reaches unity, the system will trip (blackout). The proactive cascade prediction model relies on intelligent soft computing or machine learning tools to predict cascade in advance so that the entire system can be protected against such cascade failures.

III. Theory of SVM

SVM is a kernel-based supervised computational method, which is based on the statistical learning theory [35]. An SVM kernel maps the input data points of the original input space to the higher dimensional feature space. An optimal hyperplane is determined in feature space to define a decision boundary, which separates the input data points of different classes and recognizes patterns for classification and regression.

This paper uses statistical state information of transmission lines from (3)–(8) as an input to SVM model, and SVM is trained in such a way that the direct decision function maximizes the generalization ability for the classification of grid state, normal or blackout. Consider that M, m-dimensional training inputs x_i, $(i = 1, \ldots, M)$ belong to class 1 or class 2 and then the associated labels be $y_i = 1$ for class 1 and $y_i = -1$ for class 2. If these data are linearly separable, the decision function can be determined as

$$D(x) = w^T x_i + b \tag{9}$$

where w is an m-dimensional vector, b is a bias term, and i varies from 1 to M.

$$(w^T x_i + b) \begin{cases} \leq -1, & \text{for } y_i = -1 \\ \geq 1, & \text{for } y_i = 1. \end{cases} \tag{10}$$

This equation can be expressed as

$$y_i(w^T x_i + b) \geq 1, \qquad \text{for } i = 1, 2, \ldots, M. \tag{11}$$

The optimal separating hyperplane can be obtained by solving the following convex quadratic optimization problem for w, b, and ξ [29]:

$$\min Q(w, b, \xi) = \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{M} \xi_i \tag{12}$$

subject to :

$$y_i(w^T x_i + b) \geq 1 - \xi_i,\ \xi_i \geq 0,\ \text{for } i = 1, 2, \ldots, M \tag{13}$$

where $\xi = (\xi_1, \ldots, \xi_M)^T$ is a nonnegative slack variable $(\xi \geq 0)$, and C is the margin parameter, which determines the tradeoff between the maximization of the margin and the minimization of the classification error.

To enhance linear separability, kernel trick is used in which the original input space is mapped into high-dimensional dot product space called the feature space. In SVM, according to the need of classification, a kernel has to be selected, and the values of the kernel parameter and the margin parameter C have to be determined.

The complexity and the generalization capability of the network depend on the value of the kernel parameter, since it has an influence on the smoothness of SVM response and it affects a number of support vectors. Hence, to build an optimized classifier, the optimized values of the margin parameter and the kernel parameter must be determined, which is called model selection. To avoid overfitting problem, cross-validation procedure has been used in this paper for model selection. In cross validation [36], a complete training set is divided into v

subsets of equal size; then, sequentially, one subset is tested using the classifier trained on the remaining $(v-1)$ subsets. This way, each instance of the whole training set is predicted once during SVM modeling.

IV. Proposed Proactive Blackout Prediction Model

As shown in Fig. 1, wide area monitoring system (WAMS) collects online and offline information from the grid. Online data are phase synchronized data such as voltage/current and power flow measurements of grid from phasor measurement unit (PMU). Offline data are a historical database of various operational parameters such as past blackout, islanding, or transformer outage historical data. The proposed work consists of two phases. Phase A is an offline probabilistic framework for the generation of historical database, and Phase B is a training/testing of SVM model for proactive blackouts prediction and online state classification.

1) Phase A: (Probabilistic framework for historical database) Assume a power grid having L transmission lines with maximum available capacity C (80% of IC). The lines can be characterized at time T as a state, which can be either normal (running) state or tripping (failure) state with 95% confidence level (CL, estimated range of values) and 5% level of significance (LOS, threshold of probability) [33]. Upper bound (UB) and lower bound for confidence intervals are computed from probability distribution. If the random variable (power flow data) is within the normal distribution and CDF < 1, this indicates running state (normal state), whereas CDF $= 1$ indicates blackout [20]. Prediction of the next line outage is the line that is having highest PDF value in the current sample span at time T. As shown in Fig. 1, (Phase A) blackout will occur when CDF becomes 1. The normal and blackout statistical data of a grid are stored in the historical database and further utilized as an input feature vector for training and testing of SVM.
2) Phase B: (SVM model for proactive prediction) The aim of SVM training is to train a model based on the historical data in such a way that the model can predict the target result of the test data by giving only the test data attributes. An SVM model has been developed, as shown in Fig. 1 (Phase B). Here, SVM is used as a binary classifier, and based on the output of the SVM classifier, predictions are made whether the grid is working normally or blackout will happen.

V. Case Study for Validation of the Proposed Model

The IEEE 30-bus test-bed system [37] has been built in PowerWord Simulator and used in this paper as a prototype power network. As shown in Fig. 2, the test bed consists of 30 buses, six generator buses, and 21 load buses, with 41 transmission lines consisting of 289.1-MW generation and 283.4-MW load flow capacity. The system has been modeled by

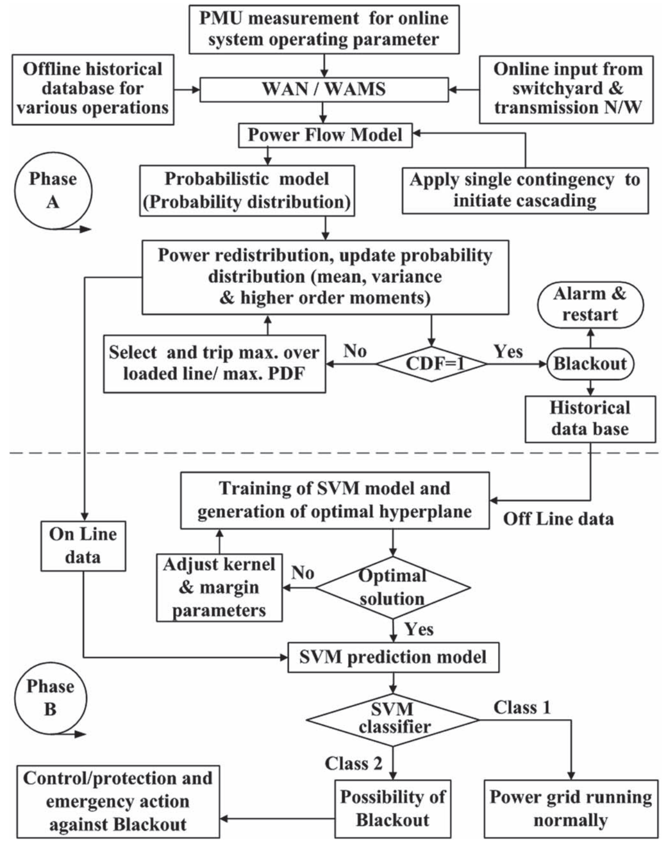

Fig. 1. Proposed model for proactive cascade failure prediction.

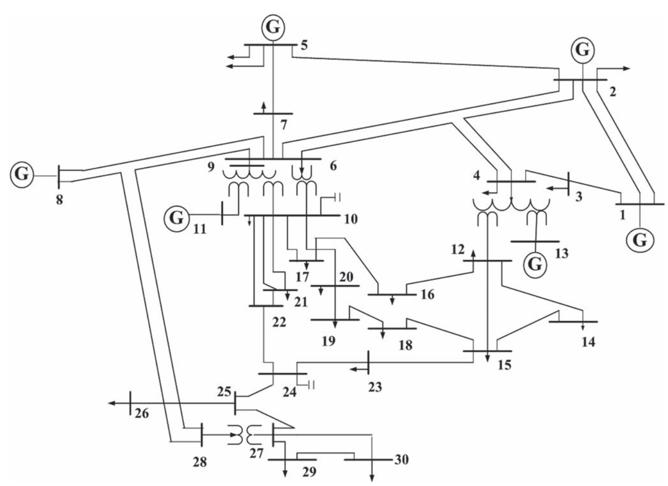

Fig. 2. One-line diagram of the IEEE 30-bus test-bed system.

maintaining impact of the line contingency as a linear function of changes in power flow on transmission line.

A. Phase A: Offline Framework for Historical Database

Cascade analysis and generation of historical database are carried out offline in Phase A.

1) Case 1, Normal working of the grid: The IEEE 30-bus test bed consists of 41 transmission lines, and at time T, it has 41 data samples of power flow, which are shown in Fig. 3. According to the central limit theorem [33],

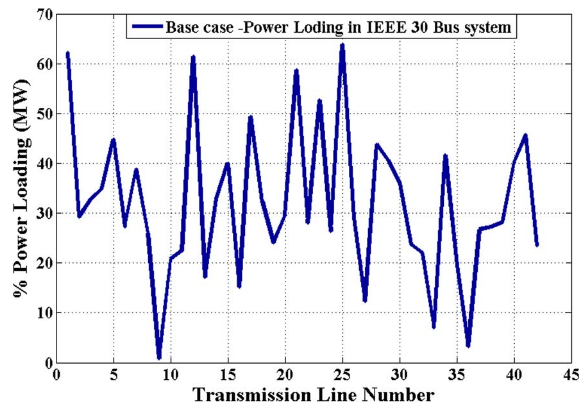

Fig. 3. Power flow in transmission line of test-bed base case.

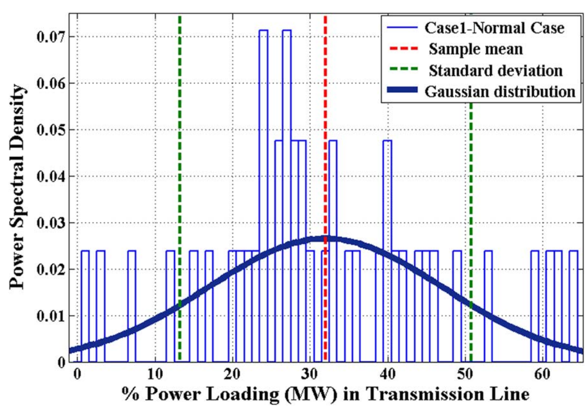

Fig. 4. Gaussian PDF as a function of power flow in the lines.

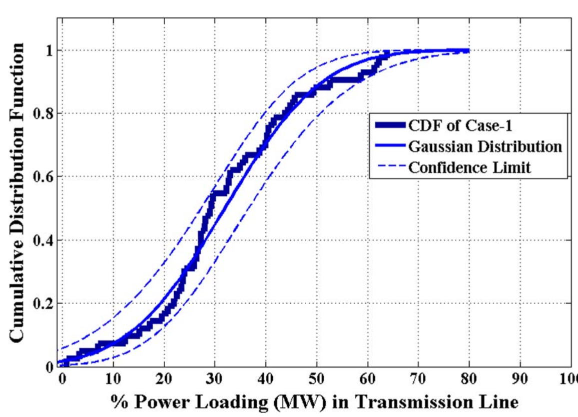

Fig. 5. CDF plot as a function of power flow in the transmission line.

the sum of a larger number (more than 30 data samples) of independent random variables tends to be a Gaussian random variable. Hence, under normal grid operation with load flow under IC, the probability distribution is considered as Gaussian distribution with 95% CL and 5% LOS, which is shown in Fig. 4. A cumulative distribution of power flow is plotted in Fig. 5, where all the data are properly fitted inside the CL, and CDF < 1; hence, all the lines are healthy at the base case. The statistical analysis of the base case is reflected in Tables I and II.

2) Case 2, Cascade failure: In Case 2, the system has been modeled with the consideration of line outage (perturbed

TABLE I
CASCADE ANALYSIS BASED ON STATISTICAL DECISION THEORY

Case	Line contingency	PDF	CDF	LB	UB	Line tripping probability
C-1	Normal grid	0.02652	0.960	0.897	0.987	–
C-2a	L21	0.02288	0.994	0.970	0.999	L22
C-2b	L22	0.01072	0.998	0.976	0.999	L29
C-2c	L29	0.00778	0.999	0.995	1	Blackout

TABLE II
HISTORICAL DATABASE PARAMETERS

Line contingency	Mean	Variance	Skewness	Kurtosis
Normal grid	μ =31.969	σ^2 =226.338	–	–
L21 Tripped	μ =33.65	σ^2 =296.862	0.5121	3.1017
L22 Tripped	μ =45.588	σ^2=1355.48	1.5751	5.4611
L29 Tripped	μ =42.576	σ^2=2048.88	2.2015	8.1761

condition). For the analysis of probability distribution under cascade failure, cascading is initiated by tripping one of the highly loaded lines, i.e., L21 between nodes 10 and 21 in Fig. 2. The probability of the next line tripping has been calculated by using statistical decision theory. Analysis of the cascading failure is carried out with Cases 2a, 2b, and 2c.

a) Case 2a is a power flow analysis after tripping of line L21. Figs. 6 and 7 are the PDF and CDF plots after line contingency, respectively. The statistical analysis of the PDF and CDF plots in Table I indicates that the next probable tripping line that is having highest PDF in the time slot T is line L22.

b) Case 2b is a contingency analysis after tripping of line L22 (between nodes 10 and 22) of the IEEE 30-bus system. As shown in Fig. 8, tripping of line L22 has changed probability distribution from Gaussian to non-Gaussian. The corresponding CDF plot is shown in Fig. 9. The nonsymmetry in Gaussian distribution is analyzed by higher order moments, and the results are given in Tables I and II. Statistical analysis and hypothesis test predict that line L29 has the highest power loading and maximum PDF and CDF values; hence, line L29 (between nodes 15 and 23) will be the next possible tripping line in Case 2b.

c) In Fig. 10, of Case 2c, tripping of L29 indicates non-Gaussian distribution with heavy tail. The corresponding CDF plot is shown in Fig. 11, where UB is crossing the limits (LOS), and the value of CDF reaches unity. The corresponding statistical analysis is shown in Tables I and II. Hence, according to statistical analysis and hypothesis testing, this scenario indicates potential cascade failure.

3) Simulation analysis: The probabilistic framework captured the essence of the cascading failure and analyzed the transition behavior of power flow in the transmission line from normal to cascade failure. This is shown in Figs. 12 and 13. As shown in Fig. 12, the behavior of the PDF changed just before the blackout (PDF of Case 2a to Case 2b), and this is the point that is used for blackout prediction in the proposed model. Similarly, a

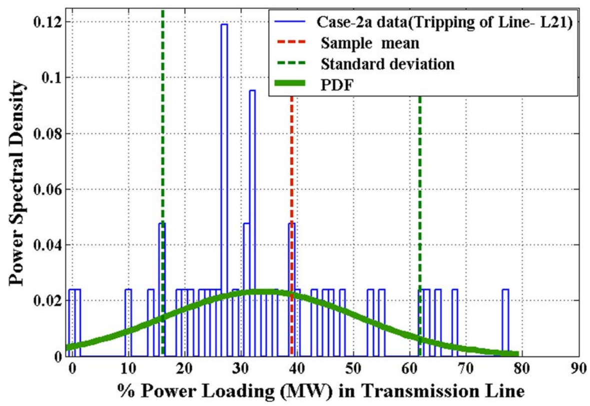

Fig. 6. Case 2a: PDF plot after tripping of line L21.

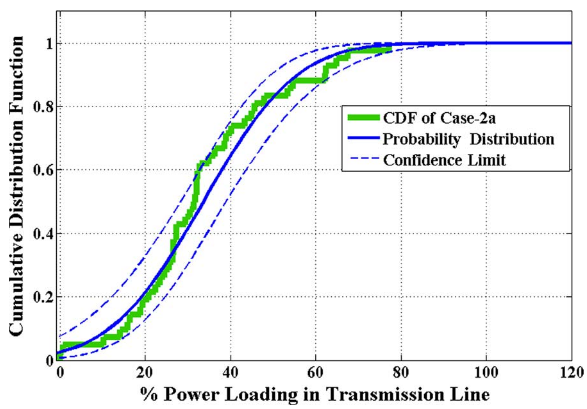

Fig. 7. Case 2a: CDF plot after tripping of line L21.

comparative CDF analysis is carried out in Fig. 13 for normal grid to blackout scenario.

The nonsymmetry in Gaussian distribution is analyzed by higher order moments using (7) and (8) and summarized in Table II. As shown in Table II, under normal condition, third- and fourth-order moments values are negligible, but as the system enters into cascading failure stage, statistical data in Table II indicate significant increase in skewness and kurtosis. This behavior highlights changes from Gaussian to non-Gaussian distribution.

The complete statistical data in Table II represent one blackout event values in historical database. Using the same procedure, a database of more than 50 such cases are simulated, verified, and stored in the historical database. Now, this historical database is used as an input feature vector to the SVM model. For training and testing of SVM, MATLAB LIBSVM toolbox is used.

B. Phase B: SVM Modeling for Proactive Blackout Prediction

Historical database of mean, variance, skewness, and kurtosis is used as input feature vectors to SVM for training, testing, and classification. SVM modeling steps are as follows.

1) Data acquisition: Historical data are acquired from Phase A of the proposed model.
2) Kernel selection: The different kernels are applied to the input data, and the results are tabulated in Table III. The result in Table III shows that the radial basis function

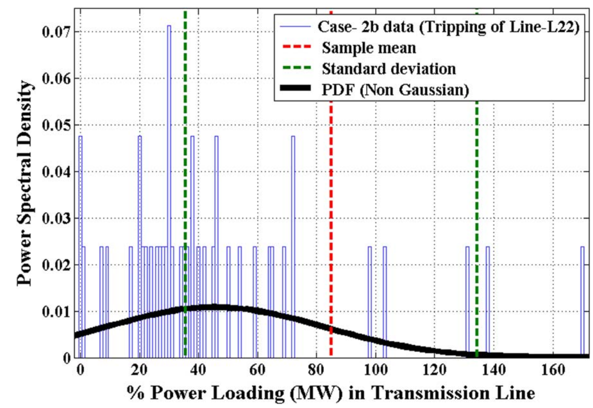

Fig. 8. Case 2b: PDF plot after tripping of line L22.

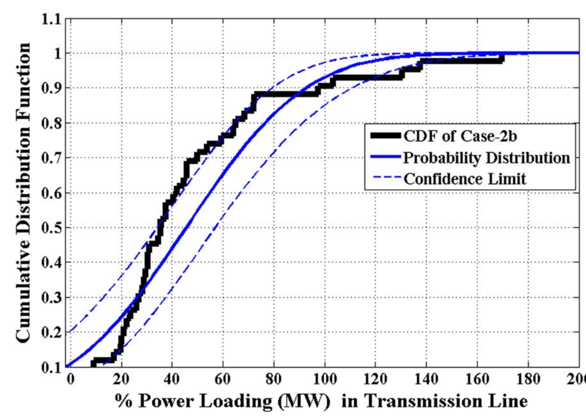

Fig. 9. Case 2b: CDF plot after tripping of line L22.

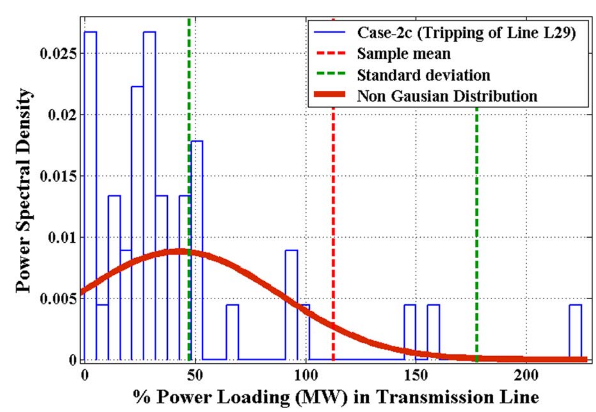

Fig. 10. Case 2c: Non-Gaussian distribution after tripping of line L29.

(RBF) kernel gives maximum training and testing accuracy. Hence, the RBF kernel has been selected.

3) Model selection: Using cross validation, find out the optimal values for kernel parameter γ and margin parameter C, which gives $C = 2$ and $\gamma = 0.25$.
4) To cross check the values of C and γ obtained from cross validation, a series of experiments have been performed. Initially, C is taken as 2, and the value of γ varies, as shown in Fig. 14. For $\gamma \geq 0.0015$, the training and testing accuracy achieved was 100%. The same procedure is followed for finding the value of C. As shown in Fig. 15,

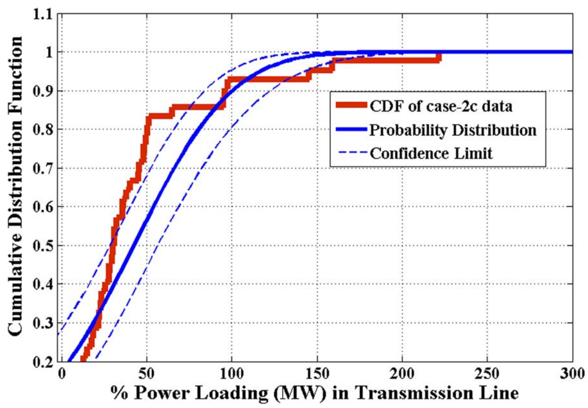

Fig. 11. Case 2c: CDF plot after tripping of line L29.

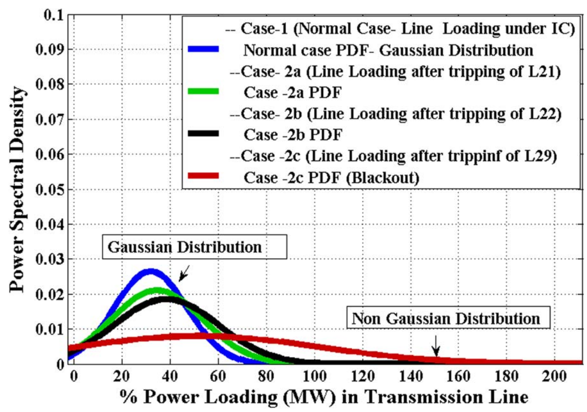

Fig. 12. Gaussian to non-Gaussian distribution of power in the test-bed system.

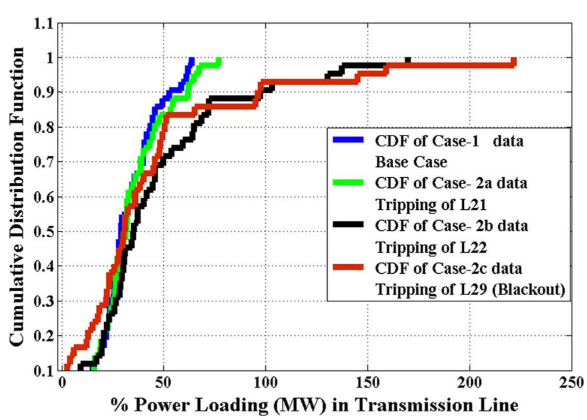

Fig. 13. CDF plots of normal to blackout cases in the IEEE 30-bus test-bed system.

TABLE III
KERNEL SELECTION FOR THE SVM MODEL

Kernel Type	Function	Training Accuracy%	Testing Accuracy%
RBF	$K(x,x^{'}) = e^{-\gamma\|\|x-x^{'}\|\|^2}$	100	100
Polynomial	$K(x,x^{'}) = (x^T x^{'} + 1)^d$	100	94.44
Sigmoid	$K(x,x^{'}) = tanh(\eta x x^{'} + \theta)$	52.63	38.89

at γ equal to 0.25, the value of $C \geq 0.73$ gives 100% training and testing accuracy. Hence, for $C \geq 0.73$ and $\gamma \geq 0.0015$, the training and testing accuracy is 100%.

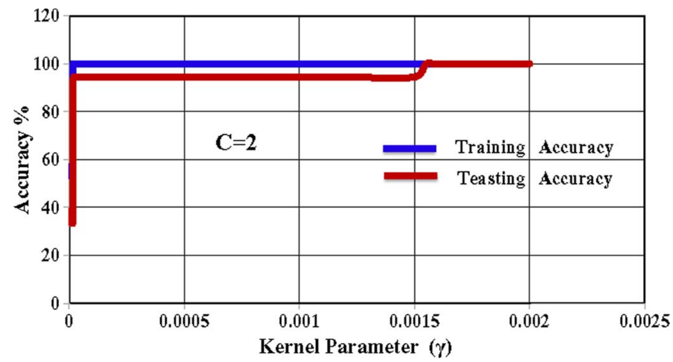

Fig. 14. SVM kernel parameter selection.

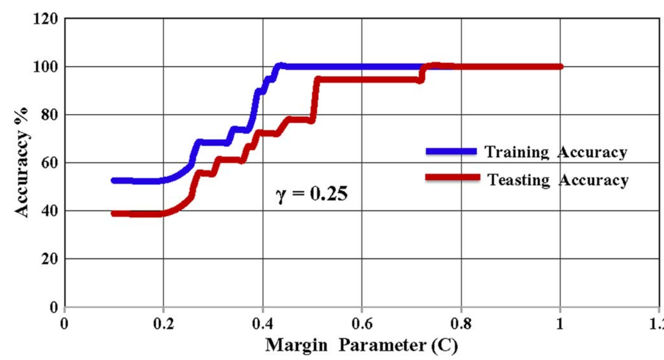

Fig. 15. SVM margin parameter selection.

VI. CONCLUSION

This paper has presented a probabilistic viewpoint of cascade failure in grid with intelligent SVM machine learning tool for proactive blackout prediction. The probabilistic framework explored the dynamics of power flow moving from pre- to post blackout. Simulation results and statistical analysis show that, under normal power flow in the grid, the probability distribution observed was Gaussian, but as cascading propagates, the probability distribution moves toward non-Gaussian distribution. The analysis of transition from Gaussian to non-Gaussian probability distribution has helped us to train the SVM model with high accuracy. The SVM output may be used to initiate emergency control systems for prevention against blackout. The proposed model can be used for proactive cascade prediction in planning and maintenance of smart grid early warning system. The future scope of the proposed model is in the field of real-world complex power networks, which demands real-time contingency analysis with self-healing and robust systems. This addresses various open research issues such as prediction and prevention of cascade failure with self-healing and robust systems. The proper integration of machine learning tools with probability theory and real-time communication technologies can achieve system-level objectives in the smart grid.

ACKNOWLEDGMENT

The authors would like to thank Dr. N. M. Singh for the technical support and guidance.

References

[1] G. Andersson *et al.*, "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1922–1928, Nov. 2005.

[2] "The enquiry committee on grid disturbance in northern region on 30th July 2012 and in northern, eastern and north-eastern region on 31st July 2012," New Delhi, India, Aug. 16, 2012. [Online]. Available: http://www.cea.nic.in/reports/articles/god/grid_disturbance_report.pdf

[3] V. Rampurkar *et al.*, "PMU based identification of low frequency oscillations: A case study," in *Proc. IEEE ISGT Asia*, Banglore, India, Nov. 10–13, 2013, pp. 1–5.

[4] F. Benzi, N. Anglani, E. Bassi, and L. Frosini, "Electricity smartmeters interfacing the households," *IEEE Trans. Ind. Electron.*, vol. 58, no. 10, pp. 4487–4494, Oct. 2011.

[5] V. C. Gungor, B. Lu, and G. P. Hancke, "Opportunities and challenges of wireless sensor networks in smart grid," *IEEE Trans. Ind. Electron.*, vol. 57, no. 10, pp. 3557–3564, Oct. 2010.

[6] T. Sauter and M. Lobashov, "End-to-End communication architecture for smart grids," *IEEE Trans. Ind. Electron.*, vol. 58, no. 4, pp. 1218–1228, Apr. 2011.

[7] Q. Yang, J. A. Barria, and T. C. Green, "Communication infrastructures for distributed control of power distribution networks," *IEEE Trans. Ind. Informat.*, vol. 7, no. 2, pp. 316–327, May 2011.

[8] G. Zhabelova and V. Vyatkin, "Multi-agent smart grid automation architecture based on IEC 61850/61499 intelligent logical nodes," *IEEE Trans. Ind. Electron.*, vol. 59, no. 5, pp. 2351–2362, May 2012.

[9] H. Li, H. Yang, F. Sun, and Y. Xia, "Sliding-mode predictive control of networked control systems under a multiple-packet transmission policy," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6234–6243, Nov. 2014.

[10] P. Siano, C. Cecati, H. Yu, and J. Kolbusz, "Real time operation of smart grids via FCN networks and optimal power flow," *IEEE Trans. Ind. Informat.*, vol. 8, no. 4, pp. 944–952, Nov. 2012.

[11] S. B. Borkowska, "Probabilistic Load Flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 752–755, May 1974.

[12] R. N. Allan, B. Borkowska, and C. H. Grigg, "Probabilistic analysis of power flows," *Proc. Inst. Elect. Eng.*, vol. 121, no. 12, pp. 1551–1556, Dec. 1974.

[13] Min, Liang, and Zhang, "Load flow with consideration of network topology uncertainties," in *Proc. 14th Int. Conf. Intell. Syst. Appl. Power Syst.*, Kaohsiung, Taiwan, Nov. 4–8, 2007, pp. 7–11.

[14] Zhang, P. Lee, and T. Stephen, "Load flow computation using the method of combined cumulants and Gram–Charlier expansion," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 676–682, Feb. 2004.

[15] A. Haghani, T. Jeinsch, and S. X. Ding, "Quality-related fault detection in industrial multimode dynamic processes," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6466–6453, Nov. 2014.

[16] Z. Wang and A. Scaglione, "A Markov-trans. model for cascading failures in power grids," in *Proc. 45th Hawaii Int. Conf. Syst. Sci.*, 2012, pp. 2115–2124.

[17] Z. Huang, C. Wang, S. Ruj, M. Stojmenovic, and A. Nayak, "Modeling cascading failures in smart power grid using interdependent complex networks and percolation theory," in *Proc. 8th IEEE Conf. ICIEA*, 2013, pp. 1023–1028.

[18] P. Hines and E. Cotilla-Sanchez, "Topological models and critical slowing down: Two approaches to power system blackout risk analysis," in *Proc. 44th HICSS*, Jan. 4–7, 2011, pp. 1–10.

[19] Rahnamay-Naeini, M. Mammoli, and A. Hayat, "A probabilistic model for the dynamics of cascading failures and blackouts in power grids," in *IEEE Power Energy Soc. Gen. Meet.*, San Diego, CA, USA, Jul. 22–26, 2012, pp. 1–8.

[20] S. R. Gupta, F. S. Kazi, S. R. Wagh, and N. M. Singh, "Probabilistic framework for evaluation of smart grid resilience of cascade failure," in *Proc. IEEE ISGT Asia*, Kuala Lumpur, Malaysia, May 20–23, 2014, pp. 255–256.

[21] C. Cecati, M. Geev, A. Piccolo, and P. Siano, "An overview on the smart grid concept," in *36th Annu. Conf. IEEE Ind. Electron. Soc.*, Glendale, AZ, USA, Nov. 7–10, 2010, pp. 3322–3327.

[22] C. Cecati and P. Siano, "Advanced computational intelligence systems for smart grids planning and management," *J. Ambient Intell. Humanized Comput.*, vol. 4, no. 6, pp. 603–604, Dec. 2013.

[23] V. N. Ghate and S. V. Dudul, "Cascade neural-network-based fault classifier for three-phase induction motor," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1555–1563, May 2011.

[24] A. Garcia-Perez, R. D. J. Romero-Troncoso, E. Cabal-Yepez, and R. A. Osornio-Rios, "The application of high-resolution spectral analysis for identifying multiple combined faults in induction motors," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 2002–2010, May 2011.

[25] T. Boukra, A. Lebaroud, and G. Clerc, "Statistical and neural-network approaches for the classification of induction machine faults using the ambiguity plane representation," *IEEE Trans. Ind. Electron.*, vol. 60, no. 9, pp. 4034–4042, Sep. 2013.

[26] F. Filippetti, G. Franceschini, C. Tassoni, and P. Vas, "AI techniques in induction machines diagnosis including the speed ripple effect," *IEEE Trans. Ind. Appl.*, vol. 34, no. 1, pp. 98–108, Jan./Feb. 1998.

[27] C. Rudin, D. Waltz, and R. N. Anderson, "Machine learning for the New York City power grid," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 328–345, Feb. 2012.

[28] Agrawal and R. Thukaram, "Fault location in power distribution system with distributed generation using support vector machine," in *Proc. IEEE PES ISGT*, Banglore, India, 2013, pp. 1–6.

[29] S. Abe, *Advances in Pattern Recognition, Support Vector Machine for Pattern Classification,* 2nd ed. London, U.K.: Springer-Verlag, 2010.

[30] Y. Zhang, "Mitigating future blackouts via smart relays: A machine learning approach," Ph.D. dissertation, Carnegie Mellon Univ., Pittsburgh, PA, USA, Jul. 2010.

[31] R. L. Sullivan and R. Lee, *Power System Planning*. New York, NY, USA: McGraw-Hill International Book Company, 1977.

[32] X.-F. Wang, Y. Song, and M. Irving, *Modern Power Systems Analysis,* 4th ed. New York, NY, USA: Springer-Verlag, Oct. 8, 2008.

[33] S. M. Ross, *Introduction to Probability and Statistics for Engineers and Scientists,* 4th ed. New York, NY, USA: Academic, 2009.

[34] A. Dwivedi, X. Yu, and P. Sokolowski, "Identifying vulnerable lines in a power network using complex network theory," in *Proc. Int. Symp. Ind. Electron.*, Seoul, Korea, Jul. 5–8, 2009, pp. 18–23.

[35] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, Jun. 1998.

[36] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," Dept. Comput. Sci., National Taiwan University, Taipei, Taiwan, Apr. 15, 2010.

[37] R. Gnanadass, Appendix—A Data for IEEE-30 Bus Test System 2011. [Online]. Available: https://www.google.com/shodhganga.inflibnet.ac.in/bitstream/10603/1221/18/18_appendix.pdf

Sudha Gupta (M'14) received the B.E. degree in electronics engineering from the Government Engineering College (Ujjain Engineering College), Ujjain, India, in 1993, and the M.E. degree in electronics engineering from Veermata Jijabai Technological Institute (VJTI), Mumbai, India, in 2001. She is currently working toward the Ph.D. degree in the field of smart grid technologies in the Electrical Engineering Department, VJTI.

She is currently an Associate Professor with K. J. Somaiya College of Engineering, Mumbai. Her research interests include probability and random variables, communication networks, image processing, and machine learning.

Ruta Kambli received the B.E. degree in electronics engineering from the University of Mumbai, Mumbai, India, and the M.Tech. degree in electrical engineering from Veermata Jijabai Technological Institute (VJTI), Mumbai.

She is currently a Senior Research Fellow with the Center of Excellence in Complex and Nonlinear Dynamical Systems, VJTI. Her research interests include cyber security in smart grids, machine learning, and biomedical engineering.

Sushama Wagh (M'09) received the M.E. degree in electrical engineering from the University of Mumbai, Mumbai, India, in 2003 and the Ph.D. degree in electrical engineering from The University of Western Australia, Crawley, Australia, in 2012.

She is currently an Assistant Professor with Veermata Jijabai Technological Institute, Mumbai. Her research interests include electrical machines, power system stability, and smart grids.

Faruk Kazi (M'12) received the Ph.D. degree from the Indian Institute of Technology Bombay, Mumbai, India, in 2009.

He is currently a Professor of electronics engineering with the Electrical Engineering Department, Veermata Jijabai Technological Institute, Mumbai. His research interests include modeling and control of complex and nonlinear dynamical systems, cyberphysical systems, and smart grids.

Prof. Kazi is a member of Working Group 10 of the Indian Smart Grid Forum.