Assessment of the Risk of Voltage Collapse in a Power System Using Intelligent Techniques

Marayati Marsadek, Azah Mohamed and Zulkifli Mohd. Nopiah

Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment, University Kebangsaan Malaysia, Bangi, 43600 Selangor, Malaysia.

Abstract: This paper describes the implementation of a fast and easy-to-use, intelligence-based algorithm to assess the risk of voltage collapse when risk is defined as the product of the event likelihood and a severity function. In the event likelihood, the effect of weather is taken into account; the failure rate of each transmission line under different weather conditions is calculated using real historical outage data. A new severity function model that utilises the voltage collapse prediction index is proposed in this paper. Two intelligent techniques, i.e., support vector machines and a generalised regression neural network are studied, and their performances are evaluated using mean absolute and mean square error. The proposed methodology has been applied in a real power system network. Simulation results show that a generalised regression neural network provides the lowest mean absolute and mean square error.

Key words: Voltage collapse, risk, GRNN, SVM.

INTRODUCTION

Recently, major blackouts throughout the world have been associated with voltage collapse (Yuri, 2005). Voltage collapse occurs for various reasons, including inadequate reactive power supply, excessive line flows or transmission line outages (Hua Wan, 2000). The restricted growth of heavily loaded power transmission networks has forced power systems to operate close to their transmission limits. As a result, power systems have become more heavily loaded and more vulnerable to disturbances. These situations have contributed to voltage instability problems. A continuous increase in load demand may slowly degrade the voltage magnitudes throughout the network; consequently, the power system is driven into a state of voltage instability, which may lead to voltage collapse. Uncontrolled voltage collapse, in turn, may lead to catastrophic system-wide collapse and blackout. To prevent such collapses from occurring, construction of a new generation of network or expansion of the current transmission network is required. Unfortunately, the initial investment cost associated with these options is not always affordable. Due to this limitation, monitoring the status of a power system under load changes and maintaining an adequate voltage profile becomes a task of critical importance.

Previously, in power system operation and planning, a power system security assessment was performed using deterministic method. In the traditional deterministic assessment practice, the power system was operated with a significant security margin that only accounts for the most credible contingencies (Hua Wan, 2000; Ning, 2006; Kirschen, 2007). Thus, this practice gives highly conservative decisions that require high-cost solutions to satisfy loading and outage conditions (Youjie Dai, 2001). In the past, the loading margin was computed using Continuation Power Flow (CPF) (Ajjarapu, 1992), L-index (Kessel, 1986), voltage collapse prediction index (VCPI) (Balamourougan, 2004) and power margin (Julian, 2000) as indicators of voltage collapse. The loading margin is an accurate index that fully accounts for power system nonlinearities and limits, such as the reactive power limit as the loading increases. The loading margin can be used with either static or dynamic system models. L-index and VCPI, whose values vary between zero and one, are determined using information obtained from load flow analysis. The status of voltage instability in a power system can be determined using the L-index and VCPI, in which zero index values represent a secure operating point. The power margin is used to track the proximity to voltage collapse; the power cannot increase if the power margin approaches zero because system collapse will occur otherwise. However, these indices alone do not provide
sufficient information regarding the condition of the current operating point. The likelihood of triggering a voltage collapse event should also be taken into account.

To allow a power system to operate closer to its limits, a more refined security assessment method is required at the planning and operating stage. A risk-based security assessment approach is one type of refined security method that takes into account the probabilistic nature of many uncertainty variables and the extent of security violations. Two important features that determine the risk index are event likelihood and event severity. To determine the event likelihood, a probability technique should be used, since the occurrence of contingencies that cause voltage collapse is stochastic in nature. To obtain the probability of these contingencies, a probability model that characterises the occurrence of contingencies is assumed, and historical data are used to determine the parameters within the assumed probability model. Within the scope of this study, only transmission line outages are considered as contingencies, and the weather is considered to be one of the factors that affect the likelihood of transmission line outage. A severity function is used to uniformly quantify the extent of voltage collapse in the power system.

Rapid growth in today’s world population and economic activities has forced power system operators to adopt a fast and reliable tool for security assessment that has the ability to provide information on the current or foreseen operating point without using a complex analytical process. This assessment should consider all contingencies, including those that have a small probability of occurrence, as their impact on the power system’s operation could be severe. Extensive studies regarding the application of a risk-based concept in voltage security can be seen in (Hua Wan, 2000; Leita da Silva, 2000; Ni, 2003; Ni, 2003; Arya, 2006; Berizzi, 2009). Only a limited set of transmission line outages are considered in (Hua Wan, 2000); thus, the risk index value computed may not provide reliable information on the current status of the power system. In (Leita da Silva, 2000), probabilistic load flow and Monte Carlo simulation has been used to determine the risk of voltage collapse without taking into account the severity of the collapse. The online procedures discussed in (Ni, 2003) and (Ni, 2003) are based on extensive power flow computations that consider the contingencies listed in the contingency list. To speed up the procedure, contingency screening is implemented by Ming Ni et al. (2003a; 2003b), but this method can be very time consuming if the power system is large. Determining the probabilistic risk of voltage collapse using the radial basis function (RBF) network implemented in (Arya, 2006) only calculates the probability of collapse. Online voltage collapse risk quantification using fuzzy techniques, as employed in (Berizzi, 2009), do not take into account the probabilistic nature of power system operation. To achieve a fast risk assessment of voltage collapse that considers the likelihood of the triggering event as well as its severity, artificial intelligence techniques that can perform parallel data processing with high accuracy and fast response times must be applied.

In this paper, the support vector machine (SVM) and generalized regression neural network (GRNN) are used to predict the risk of voltage collapse in a power system by incorporating a weather-dependent parameter selection to calculate the probability of transmission line outage. The purpose of the study is to compare the performance of these two developed intelligent techniques to choose the best tool to predict the risk of voltage collapse in a power system. To validate the proposed intelligent techniques, they are implemented on a real practical power system.

**Risk of Voltage Collapse:**

Risk-based techniques have led to a paradigm shift in security assessment. The risk of voltage collapse (VC) is defined as follows:

\[
RISK(VC) = \sum_{i=1}^{N} \text{Prob}_i(E) \times \text{Sev}_i(E)
\]

where \(\text{Prob}(E)\) is the probability of the event that triggers voltage collapse, \(\text{Sev}(E)\) is the event’s severity and \(N\) is the number of contingencies considered in the study.

Contingencies such as transmission line or generator outages may result in voltage instability in the power system. The power system is secure if none of the contingencies causes a voltage collapse. In this study, only uncertainty in the transmission line outage is considered.

**Probability of Transmission Line Outage:**

The probability of a transmission line outage that can cause a security violation is referred to as the event
likelihood, and it is assumed to have the characteristics of a Poisson process (Ni, 2003; Ni, 2003; Arya, 2006). A Poisson process counts the arrival of events within a pre-defined time interval. The probability model of a Poisson random variable describes a phenomenon that occurs randomly in time, where the time of each occurrence is completely random and the average number of occurrences per unit time is known. In a transmission network consisting of a total of N lines, the probability for ‘N-1’ contingencies can be written as (Marsadek, 2011).

\[ \text{Prob}_j (E) = (1 - e^{-\lambda t}) \times e^{-\sum_{j \neq i} \lambda_j} \]  

(2)

where \( \lambda \) is the failure rate per unit time.

In this application, the parameters that must be determined are the failure rates of each transmission line, which can be calculated from the historical outage data. The historical outage data obtained from a utility for a six-year period (2002-2007) show that for many transmission lines, little or no outage history exists. Transmission line outage is a weather-dependent event, and therefore, the failure rate varies with the weather conditions.

The data pooling procedure described in (Fei Xiao, 2006) is applied to address the above-mentioned issues. Given a set of historical data, the estimator is improved by grouping the lines and pooling their outage history data so that data from different components having similar characteristics are combined. As a result, the size of the database in which the probability model parameter is computed increases and a better estimator of the failure rate of a component is obtained.

**Table 1:** Data pooled to estimate the failure rate

<table>
<thead>
<tr>
<th>Pool</th>
<th>Pool 1</th>
<th>.....</th>
<th>Pool Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather</td>
<td>W_1</td>
<td>.....</td>
<td>W_M</td>
</tr>
<tr>
<td>Line</td>
<td>outage</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

First, the transmission outage data are pooled by geographical area under the assumption that the weather is relatively homogenous throughout the region. Second, the outage data is rearranged according to its pool and weather condition, as shown in Table 1. In Table 1, transmission lines are grouped into ‘Q’ pools and each outage in the respective pools is classified as being associated with one of ‘M’ weather conditions. Third, the failure rate per unit mile of a transmission line in each data pool associated with a specific weather condition is calculated as

\[ \lambda_{km}^{(n)} = \frac{n}{l \times T} \]

(3)

where \( k \) is defined as the pool number, \( m \) is the weather condition, \( n \) is the number of observed failures, \( l \) is the total length of the line experiencing the outage in pool ‘\( k \)’ under the weather condition ‘\( m \)’ and \( T \) is the sample size.

The flowchart of the proposed algorithm to estimate the failure rate is depicted in Fig. 1.
A severity function is used to uniformly quantify the severity of network performance. In their paper, Ni et al. (2003) highlighted the criteria for a good severity function. In this study, a severity function based on the value of VCPI is selected because the consequences are easily understood in terms of network parameters, and it can reflect the relative severity of different problems, which enables a composite risk index to be calculated.

VCPI is calculated at every bus using the voltage phasor measurement and network admittance matrix. It is given by the following relationship (Balamourougan, 2004):

$$\text{VCPI}_{k_{	ext{bus}}} = \left| 1 - \frac{\sum_{m=k}^{M} V'_m / V_k}{\sum_{m=k}^{M} V'_m / V_k} \right|$$  

where $V_k$ is the voltage phasor at bus ‘$k$’, $V'_m$ is the voltage phasor at bus ‘$m$’, $Y_{km}$ is the admittance between bus ‘$k$’ and ‘$m$’ and $M$ is the number of the bus in the network.

If the VCPI value is zero, the bus is considered voltage stable, and if the VCPI value is one, a voltage collapse is said to occur. Figure 2 shows the variation of VCPI value with respect to the load margin.

From Figure 2, it is noted that the VCPI value is zero when no load is assigned to a bus in which at this condition the load margin is maximum. The VCPI value increases from zero to one as percentage difference between the loadability and load decreases from 100% to 0%. The bus with the largest VCPI value is the weakest bus in the network. Therefore the severity function due to transmission line outage leading to voltage collapse utilising the VCPI index is defined as follows,
Fig. 2: VCPI with respect to load margin.

\[ \text{SEV} (E) = \max (VCPI_1, VCPI_2, \ldots, VCPI_M) \] (6)

where \( M \) is the number of buses in the power system network.

**SVM and GRNN Theory:**

In this section, the background theory of SVM and GRNN are described briefly.

**Support Vector Machine:**

SVM, which was first introduced by Vapnik in the late 1960s, is an excellent tool for classification and regression problems of good generalisation performance. The SVM, which is based on statistical learning theory, is a general classification method (Vapnik, 1998; Cristianini, 2000). SVM works in the high-dimensional feature space formed by the nonlinear mapping of an \( N \)-dimensional input vector \( x \) into a \( K \)-dimensional feature space \( (K > N) \) through the use of the function \( \phi(x) \). \( K \) is the number of hidden units, which are also known as the support vectors.

Given a set of data points \( \{(x_1, y_1), \ldots, (x_i, y_i)\} \) such that \( x_i \in R^N \) is an input and \( y_i \in R^N \) is a target output, the estimated function is as follows:

\[ f(x) = w^T \phi(x) + b \] (7)

Learning is motivated through a linear \( \varepsilon \)-insensitive loss function, which is defined as

\[ L(x, y, f) = \max (0, |y - f(x)| - \varepsilon) \] (8)

where \( f(x) \) is the output of the SVM network and \( \varepsilon \) is the defined accuracy.

The learning task in SVM is defined as the minimisation of the sum of the linear \( \varepsilon \)-insensitive losses, which are given as

\[ \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{N} L(x_i, y_i, f) \] (9)

where \( w = [w_0, w_1, \ldots, w_K]^T \) is the weight vector matrix and \( \| w \|^2 \) is less than a user-specified constant \( C \).

The standard form of a SVM for solving regression problems is as follows:

\[ \min_{w,b,\xi^i, \xi^{i'}} \left\{ \frac{1}{2} w^T w + C \sum_{i=1}^{N} (\xi_i + \xi_i') \right\} \] (10)

where \( \xi_i \) and \( \xi_i' \) are the non-negative slack variables.

Minimisation of the cost function specified in (10) is subjected to the following boundary constraints:
To solve the constrained optimisation problem, a Lagrangian function and multipliers are employed. The minimisation of the Lagrangian function is transformed into a dual problem, which is defined as

\[
\max \left\{ \sum_{i=1}^{d} y_i (\alpha_i - \alpha'_i) - \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} (\alpha_i - \alpha'_i)(\alpha_j - \alpha'_j)K(x_i, x_j) \right\}
\]

where

\[K(x_i, x_j) = \varphi^T(x_i) \varphi(x_j)\]

The function described in (18) is subject to the following constraints:

\[
\sum_{i=1}^{d} (\alpha_i - \alpha'_i) = 0
\]

\[
0 \leq \alpha_i \leq C
\]

\[
0 \leq \alpha'_i \leq C
\]

The output of an SVM network can then be expressed as

\[
f(x) = \sum_{i=1}^{d} (\alpha_i - \alpha'_i)K(x_i, x) + b
\]

where \(b\) is the bias term and the term \(K(x_i, x_j)\) is the kernel function.

In general, it is difficult to determine the type of kernel function to use for specific data patterns. However, any function that satisfies Mercer’s condition, as suggested by Vapnik (Cristianini, 2000), can be used as a kernel function. Kernels are selected based on the data structure and the type of boundaries between the classes. In this work, the radial basis function (RBF) kernel is used, and it is defined as

\[
K(x_i, x_j) = \exp \left( -\frac{||x_i - x_j||^2}{\sigma^2} \right)
\]

where \(\sigma\) is a parameter that controls the width of the RBF kernel function.

The selection of the model parameters \(\sigma^2\) and \(C\), is important to ensure the SVM accuracy when solving the regression problem. In this work, LIBSVM (Chih-Chung Chang, 2001) is used as a tool to create the SVM, and the parameter selection is optimised using the cross validation method.

**Generalized Regression Neural Network:**

A GRNN has the ability to estimate a function directly from a given training data. The network structure of a GRNN as shown in Fig. 3 comprises two layers, namely, the radial basis and special linear layers. The input neurons are made up of \(p\) neurons and a bias vector \(b\), where \(p\) is the dimension of the input vector \(x\) and \(b\) is a column vector whose elements are set to 0.8326/SPREAD. SPREAD is a user-defined parameter determined by trial and error that measures the distance an input vector must be from a neuron’s weight vector using a value of 0.5 (Vapnik, 1998). All input units are fully connected to the neurons in the radial basis layer.
The input neurons receive the input vector and produce a distance vector between the input and its weight vector, \( w \), using the Euclidean distance weight function, \( ||d|| \). Then, the distance vector is passed through the radial basis layer and adjusted by the bias. The radial basis layer is made up of a radial basis neuron whose transfer function is shown in Fig. 4. The mathematical representation of the radial basis function is

\[
a_1 = \exp \left(-n_1^2\right)
\]  

where \( n_1 \) is the output from the input units.

A pattern neuron is used to combine and process the data systematically such that the relationship between the input and the proper response is "memorised". The radial basis layer receives input and produces output that has a maximum of one when the input is zero. A dot product operation is then performed between the output of the radial basis layer and the weight vector \( LW_{2,1} \), where \( LW_{2,1} \) is equal to the target vector. The output of the dot product operation is then passed on to the special linear layer, whose transfer function is shown in Fig. 5. From the figure, \( n_2 \) is the output of the radial basis layer, or in this case, \( a_1 \cdot f(x) \) can be written in the following form:

\[
f(x) = \text{purelin}(n_2)
\]
**Development of the SVM and GRNN Models:**

To demonstrate the effectiveness of the proposed techniques, they have been applied to a real power transmission network at a voltage level of 275 kV that consists of 87 buses. The network includes 57 single, 54 double and 3 quadruple lines. The single line diagram of the test system is shown in Fig. 6.

![Linear transfer function](image)

**Fig. 5:** Linear transfer function.

The test system is divided into four regions, namely, Northern (N), Eastern (E), Central (C) and Southern (S) Regions with a total real and reactive load power at base case condition of 10920 MW + j2420 MVar.

![The test system](image)

**Fig. 6:** The test system.
**Estimation of the Failure Rate:**

In Section 2.1, a data pooling procedure is proposed to estimate the Poisson parameters, which, in this case, are the failure rates of each transmission line. Assuming that the weather in each region is homogeneous, data pooling is performed based on region. Table 2 shows the transmission line pools, their respective regions and the number of transmission lines in each pool. Pools 5 through 9 represent the grouped transmission lines that connect two regions.

**Table 2: Transmission line pool.**

<table>
<thead>
<tr>
<th>Pool</th>
<th>Region</th>
<th>No. of transmission line</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Northern</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>Central</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>Southern</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>Eastern</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Northern - Central</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Central – Southern</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Eastern – Northern</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Central – Eastern</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>Eastern - Southern</td>
<td>1</td>
</tr>
</tbody>
</table>

From the historical outage data obtained from January 2002 to December 2007, the weather condition can be classified as one of three types, namely, clear, cloudy and rain. Using (3), the failure rate per unit mile per year for each respective pool is calculated. A lookup table consisting of length in miles \( l \) and pool number \( k \) for each transmission line has been developed to compute the failure rate for each line. The failure rate for each transmission line with respect to weather condition is calculated by multiplying the failure rate per unit mile per year by the line length.

**Selection of Input and Output Features:**

In the neural network implementation, the number of input features greatly affects the training process and accuracy. A previous study (Marayati Marsadek, 2010) has shown that the severity of each contingency is directly associated with the total load demand. The probability of each contingency occurring is calculated using the failure rate of each transmission line and it is directly linked to the weather condition. To develop the SVM and GRNN, real data regarding the load assigned to each load bus and the weather condition experienced by each transmission line are selected as input variables, and the risk of voltage collapse is the output variable. Numerical values of 1, 2 and 3 are assigned to represent clear, cloudy and rainy weather, respectively. Table 3 summarises the selected input features; the total number of input features is 161.

**Table 3: Selected input features for SVM and GRNN.**

<table>
<thead>
<tr>
<th>No.</th>
<th>Feature description</th>
<th>Number of features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Load assigned to each load bus (MVA)</td>
<td>54</td>
</tr>
<tr>
<td>2</td>
<td>Weather Condition experienced by each transmission line</td>
<td>107</td>
</tr>
</tbody>
</table>

**Data Preparation:**

Sufficiently large numbers of training samples have been generated prior to the implementation of the intelligent techniques to ensure good generalisation. The load at each load bus is randomly changed between the base case and 130% of the base case load. A total of 1100 training instances are generated in which 733 samples are used for training purposes while the remaining 367 samples are used for testing. During the training stage, training and testing data sets are randomly selected. The SVM and GRNN are trained separately.

To determine the value of the severity function, the voltage magnitudes of all buses subjected to transmission line outages must be examined. To simulate difference contingencies, the Power System Analysis Toolbox (PSAT) (Federico Milano,) is used to perform the power flow simulations. A contingency list that consists of 107 single line outages has been considered. All contingencies have been simulated, and their corresponding severity index values are calculated for each operating point. The outage probability for each transmission line is calculated using the value of the failure rate determined from the data pooling method.

**Data Pre-processing:**

In order to have a more accurate prediction result, the input and target data needs to be represented in a normalised form that ranges between zero and one. This prevents any parameter from dominating the output.
value and also provides better convergence and accuracy of the learning process. The normalisation procedure is as follows:

\[ d_k = \frac{d_k^\prime - d_{k, \text{min}}}{d_{k, \text{max}} - d_{k, \text{min}}} \]

where \( d_k^\prime \), \( d_k \), \( d_{k, \text{min}} \), and \( d_{k, \text{max}} \) represent the normalised value, raw input or target value, minimum and maximum values of the respective features, respectively.

**Performance Measures:**

To evaluate the effectiveness of the SVM and GRNN, the Mean Absolute Error (MAE) and Mean Square Error (MSE) are considered and written as

\[
\text{MAE} = \frac{1}{t} \sum_{q=1}^{t} |y_q - f_q(x)|
\]

\[
\text{MSE} = \frac{1}{t} \sum_{q=1}^{t} (y_q - f_q(x))^2
\]

where \( t \) is the number of the training sample and \( y \) and \( f(x) \) are the real and predicted values, respectively. MAE is a performance measure that indicates how close predictions are to the actual outcomes. The assessment of the quality of an estimator in terms of its variance and unbias is given by the MSE.

**RESULTS AND DISCUSSION**

The SVM and GRNN intended for prediction of risk of voltage collapse are implemented on a real power system. In order to determine the likelihood of the occurrence of a security violation, the failure rate for each line with respect to the weather condition is first computed. Using the data pooling procedure explained in Sections 2.1 and 4.1, each transmission line is assigned to a pool. In this study, pooling is performed based on the geographical location of the transmission line. A real historical outage data set obtained from the utility is used to compute the failure rate of each transmission line. Table 4 shows the failure rate associated with the weather conditions computed using the data pooling method for each pool.

<table>
<thead>
<tr>
<th>Pool</th>
<th>Failure rate (failure/ mile / year)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clear</td>
</tr>
<tr>
<td>1</td>
<td>2.69E-03</td>
</tr>
<tr>
<td>2</td>
<td>1.50E-03</td>
</tr>
<tr>
<td>3</td>
<td>3.44E-03</td>
</tr>
<tr>
<td>4</td>
<td>1.09E-03</td>
</tr>
<tr>
<td>5</td>
<td>1.65E-03</td>
</tr>
<tr>
<td>6</td>
<td>9.11E-03</td>
</tr>
<tr>
<td>7</td>
<td>1.67E-03</td>
</tr>
<tr>
<td>8</td>
<td>1.44E-03</td>
</tr>
<tr>
<td>9</td>
<td>8.69E-04</td>
</tr>
</tbody>
</table>

From the results presented in Table 4, it is apparent that transmission line outage is a rare event for which the occurrence rate is very low. The effect of weather on the occurrence of transmission line outage can be seen in pools 1, 2, 3 and 4. For example, in Pool 1, a transmission line outage is more likely to occur on a rainy day because the failure rate of the transmission line in Pool 1 during a rainy day is higher than the failure rates associated with clear and cloudy days. However, the variation of failure rate with respect to weather condition is not revealed in pools 5, 6, 7, 8 and 9. The constant failure rate in these 5 pools may be due to the fact that few transmission lines are located in these respective pools, and therefore, few historical outages occurred within the specified period.
To ensure that the SVM and GRNN can achieve the best generalisation ability, user-defined parameters have to be first optimised. For parameter tuning, the cross validation method from the LIBSVM software package is used. To achieve good generalisation accuracy in the implementation of the GRNN, the SPREAD is adjusted heuristically. In order to select the optimum parameters for the SVM and GRNN, training and testing processes are carried out using different parameters values, and their performances are evaluated using MAE and MSE. The values of the user-defined parameters used in LIBSVM and the GRNN to predict the risk of voltage collapse and its respective performance measures are summarised in Tables 5 and 6, respectively.

### Table 5: User-defined parameters for LIBSVM

<table>
<thead>
<tr>
<th>C</th>
<th>$s^2$</th>
<th>MAE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>0.049</td>
<td>0.0035</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.0967</td>
<td>0.0127</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0.0418</td>
<td>0.0027</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0.0431</td>
<td>0.0028</td>
</tr>
</tbody>
</table>

### Table 6: User-defined parameters for GRNN

<table>
<thead>
<tr>
<th>SPREAD</th>
<th>MAE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.015</td>
<td>0.000945</td>
</tr>
<tr>
<td>0.15</td>
<td>0.0253</td>
<td>0.0025</td>
</tr>
<tr>
<td>0.20</td>
<td>0.0216</td>
<td>0.0018</td>
</tr>
<tr>
<td>0.30</td>
<td>0.0291</td>
<td>0.0033</td>
</tr>
</tbody>
</table>

It can be seen from Table 5 that the minimum MAE and MSE in the predicted risk of voltage collapse using LIBSVM are obtained when the user defined parameters $C$ and $s^2$ are tuned to 3 and 0, respectively. The evaluation of the prediction performance of GRNN in risk assessment of voltage collapse shown in Table 6 indicates that the best accuracy is obtained when SPREAD is adjusted to 0.10. From the simulation result, it is found that the optimum SVM model required to assess the risk of voltage collapse contains a total of 21 support vectors. The GRNN network developed to predict the risk of voltage collapse contains 161 neurons in the input layer, 733 neurons in the second layer and 1 neuron in the output layer.

The performances of SVM and GRNN are compared to select the best prediction model of the risk of voltage collapse. The corresponding performance measure given in Table 7 show that the prediction obtained using GRNN is the most satisfactory because it provides the lowest MAE and MSE.

### Table 7: Optimal user-defined parameters

<table>
<thead>
<tr>
<th>Method</th>
<th>Optimal User Defined Parameter</th>
<th>MAE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>$C = 3$, $s^2 = 0$</td>
<td>0.0418</td>
<td>0.0027</td>
</tr>
<tr>
<td>GRNN</td>
<td>SPREAD = 0.10</td>
<td>0.015</td>
<td>0.000945</td>
</tr>
</tbody>
</table>

Fig. 7: Absolute error.
In order to examine the performance of the developed intelligent model in more detail, the absolute error curve obtained during the testing stage is shown in Fig. 7. As can be seen in Fig. 7, in most cases, the absolute error of each testing instance executed by the SVM model is much higher than those associated with GRNN. Furthermore, the GRNN model is also able to predict the risk of voltage collapse with zero absolute error. Based on the results, to achieve optimal performance of the prediction tool used to assess voltage collapse in a test system, it is appropriate to employ a GRNN to predict the risk of voltage collapse.

Conclusion:
An intelligent tool that predicts the risk of voltage collapse in a power system that accounts for the effect of weather on the likelihood of the occurrence of a security violation has been presented. Two neural network models, SVM and GRNN, are studied, and their performances are evaluated using MAE and RMSE to select the best neural network model for risk prediction. The optimal risk prediction model is chosen from the candidate model with a minimum performance measure in which the test results showed that GRNN with an optimised user-defined parameter is the most accurate prediction model for voltage collapse when compared to SVM. The advantage of this model is that less information is needed to determine the risk index value, and thus, it can be used as an online tool to monitor the risk. Further related work may apply feature extraction to the proposed methodology to reduce the number of input features.

REFERENCES