Machine learning for frequency estimation of power systems

E.S. Karapidakis *

Company for Support and Development of Cretan Enterprises, 50 Giamalaki 50 Str., Heraklion, 71202 Crete, Greece

Received 26 January 2004; received in revised form 8 March 2005; accepted 1 April 2005

Abstract

In this paper the application of machine learning techniques for on-line dynamic security assessment of power systems is presented. Decision trees (DT), artificial neural networks (ANN) and entropy networks (EN) are developed and applied on the power system of Crete, the largest Greek island. Comparison of these methods reveals their relative advantages and disadvantages. These methods have been integrated in the dynamic security assessment module of the advanced control system of Crete island, helping to identify the operating conditions and parameters that lead to a less robust operation of the system. The results are considered very satisfactory, both in accuracy that increases the reliability of the method and in computational time, which is a necessity for real time applications.

© 2005 Published by Elsevier B.V.

Keywords: Power systems; Dynamic security assessment; Machine learning; Decision trees; Entropy trees; Neural networks and energy management systems

1. Introduction

The recent developments in electric energy systems, i.e. the on-going liberalization of the energy markets, the pressing demands for power system efficiency and power quality, the increase of dispersed generation and power exchanges among utilities, dictate the needs for improvements in the power system planning, operation and control. Machine learning techniques together with the traditional analytical techniques can significantly contribute in the solution of the related problems.

This paper focuses on the problems of autonomous power systems, like the ones operating in large islands. These systems, cause to non-support of any interconnection, face increased problems with respect to their operation and control, i.e. high frequency and voltage instability. A reliably operating autonomous power system should continuously supply power in an economic way, while the frequency of the system should be maintained in permissible limits. The most economic operation of the system is a necessity into the new framework of “open” energy market. In addition, there are many technical constraints, which must be considered as they may seriously reduce the system operating performance. A better system performance can be achieved, if the system is controlled from an advanced energy management system (EMS) [1,2].

* Tel.: +30 6946300081; fax: +30 28210230003.
E-mail address: mkarapidakis@ank.gr.
One of the basic functions of such an EMS is online security assessment. The application of machine learning techniques requires three general steps, namely application of: (1) random sampling techniques, to scan all relevant operating states of the system and simulation of these states under various pre-determined disturbances to build the knowledge base (KB), (2) various techniques to extract and synthesize relevant information and to reformulate it in a suitable way for decision-making and (3) application of the extracted synthetic information, either in real time, for fast and effective decision making, or in off-line study environment, so as to gain new physical insight and to derive better system operation and planning strategies [3].

In our application the on-line dynamic security assessment of Crete, the largest Greek island, has been investigated. Following a repeated procedure a number of initial operation points (OPs) have been calculated by varying the load demand and the production of existing renewable sources and various disturbances were simulated running EUROSTAG. Decision trees (DT) are first developed and applied in order to identify the most important system parameters [4]. Artificial neural networks have been trained using the same learning sets as DTs and their results are evaluated [5,6]. Finally, entropy networks, a hybrid neural network approach is applied aiming at combining the attractive features of the two techniques, namely the simplicity and transparency of decision trees and the information accuracy of multi layer perceptrons [7]. Combination of previous structures and analytical mathematical methods lead to a more robust operation of the system, while application of these techniques in the power system of Crete and comparison with actual events has shown that security can be assessed accurately online [8,9].

The paper describes the application of automatic learning techniques for dynamic security evaluation of isolated power systems. These techniques have been integrated in a module of the advanced control system of Crete island, helping to identify the operating conditions that lead to a less robust operation of the system. The results are considered highly satisfactory, both in accuracy that increases the reliability of the method and in computational time, which is a necessity for real time applications.

2. Machine learning in power systems

The study case system is a realistic model of the power system of Crete. The power system of the island of Crete is the largest autonomous power system in Greece with the highest rate of increase in energy and power demand nationwide. The transmission network consists mainly of 150 kV lines and some 66 kV lines. The generation system consists of two major power plants one in Linoperamata and one in Chania located near the major load points of the island. There are 18 thermal, oil-fired generating units with a total capacity of 523.2 MW installed, including 6 steam units of total capacity 112 MW, 4 diesel units with 49.2 MW, 7 gas turbines with 228.6 MW and one combined cycle plant with 133.4 MW. One characteristic of the load profile is the significant difference between low load and peak load recorded values, both within a day and within a year. The annual peak load demand occurs on a winter day and overnight loads can be assumed to be approximately equal to 25% of the corresponding daily peak loads. Normally, the peak load exceeds 350 MW, while the lowest load is about 100 MW. The combined cycle and the steam units mainly supply the base-load. The gas turbine units normally supply the peak load at a high running cost that increases significantly the average cost of the electricity being supplied. Eleven wind parks of a total capacity of 80 MW have already been installed and many more are planned to be installed in Crete in the near future, operated mainly by independent power producers [10].

The application of automatic learning techniques is based on previous knowledge about the behavior of the system, obtained from a large number of off-line dynamic simulations that define a data set. This data set is split into a learning set (LS), used to derive security evaluation structures, and a test set (TS) used for testing the developed structures. The data set consists of a large number of operating points (OPs) each characterized by a vector of pre-disturbance steady-state variables, called attributes [8]. These can be directly measured (powers, voltages, etc.) or indirectly calculated quantities (wind penetration, spinning reserve, etc.). For the creation of the data set, the initial OPs are obtained by varying randomly the load for each load busbar and the wind power for each wind park. These variables are assumed to follow normal distributions around four starting operating profiles, as it is shown in Fig. 1.
The total load demand follows a ±30 MW change around the four above operating profiles (final load range: 90–330 MW). A dispatch algorithm approximating actual operating practices followed in the control system of Crete is applied next in order to complete the pre-disturbance OPs. This means that the steam units and the combined cycle plant cover the base load, while the gas turbines mainly supply the peak. For each of the OPs, a number of disturbances have been simulated using EUROSTAG [11]. Two major disturbances have been finally selected after studying extensively the behavior of the network. These are:

- outage of a major gas turbine and
- three phase short circuit at a Hbus-bar.

In fact, a unit disconnection is a frequent event and a three-phase fault, although rare, is a severe event that can occur during stormy conditions. For each OP, the minimum value of system frequency and the maximum rate of frequency change are recorded. Both of these parameters are checked against the values that activate the under-frequency relays of the system, and the OPs are then labeled as secure/insecure. Finally, the selected security criterion that exploits the minimum frequency of the system after the disturbance follows the next rule:

\[
\text{If } f_{\text{min}} > 49 \text{ Hz and } \text{df/dt} < 0.4 \text{ Hz/s then the OP is secure}
\]
\[
\text{else it is insecure}
\]

The list of the 10 final selected attributes that characterize each OP, includes namely:

- Active power of first unit group (combine cycle) and its spinning reserve, Comb and CombRes, respectively.
- Active power of second unit group (diesel units) and its spinning reserve, Diesel and DieselRes, respectively.
- Active power of third unit group (steam units) and its spinning reserve, NStm and NStmRes, respectively.
- Active power of fourth unit group (gas units) and its spinning reserve, OStm and OStmRes, respectively.
- Active power of fifth unit group (gas units) and its spinning reserve, NGas and NGasRes, respectively.

Using previous described approach 5735 acceptable OPs have been obtained, which were divided into the LS comprising 3748 OPs and the TS comprising 1987 OPs. For each OP the maximum frequency deviation and the rate of change of frequency are recorded.

3. DTs implementation

The decision tree methodology is a non-parametric learning technique able to produce classifiers about a given problem in order to deduce information for new unobserved cases. The construction of a DT starts at the root node with the whole LS of pre-classified OPs. These OPs are analyzed in order to select the test-\(T\) that splits them “optimally” into a number of most “purified” subsets. For the sake of simplicity, a two-class (safe–unsafe) partition is considered. The test-\(T\) is defined as

\[
T : A_i \leq t
\]

where \(t\) is the optimal threshold value of the chosen attribute \(A_i\). The selection of the optimal test and optimal threshold is based on maximizing the additional information gained through the test. The selected test is applied to the LS of the node splitting it into two subsets, corresponding to the two successor
nodes. The optimal splitting rule is applied recursively to build the corresponding sub-trees. In order to detect if one node is terminal, i.e. sufficiently high percent of only one class, the stop splitting rule is used, which checks whether the entropy of the node is lower than a present minimum value. If it is, the node is declared a leaf, otherwise a test-\( T \) is sought to further split the node. If the node cannot be further split in statistically
significant way, it is termed a deadend, carrying the two class probabilities estimated on the basis of the corresponding OPs subset [3,4].

Based on the previous learning set, two DTs have been developed, each corresponding to one of the disturbances simulated. In the root node and in the non-terminal nodes, information related with the number of OPs, the safety index and the splitting test are included. In the terminal nodes information about the number of OPs that belong there, the safety index of those OPs and the type of the node are given. The developed DTs for both disturbances are shown in Figs. 2 and 3 correspondingly, while the classification performance evaluation for both DTs is shown in Tables 1 and 2.

4. ANNs implementation

Artificial neural networks, as decision trees, are based on previous knowledge about the behavior of the system obtained from a large number of off-line dynamic simulations. Neural Network Toolbox of Matlab Package was used for this study [12]. After analytical studies, two ANNs have been selected for both disturbances. For each case, we use the same ANN architecture that is shown in Fig. 4, but different weight matrixes and bias vectors.

The network shown in Fig. 4 is a fully connected feed-forward neural network with two layers [5]. Each layer has a weight matrix $W$, a bias vector $b$ and an output vector that is input for the next layer. It has five inputs ($I_i$), one hidden layer with four neurons and two outputs ($O_i$). The total number of weights is 28 and of biases is 6.

Feed-forward networks often have one or more hidden layers of sigmoid or tan-sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range $-1$ to $+1$. The transfer function of the hidden layer, in our networks, is the Tan-Sigmoid function that generates outputs values between $-1$ and $+1$. The transfer function of the output layer is the linear function in order to give the appropriate range of output values for our study.

After the construction of the ANNs as it is described in previous paragraph, the networks are trained with constant learning rate of 0.1, momentum of 0.06 and maximum error tolerance of 0.001. Back propagation algorithm was used for learning the study ANNs. The back propagation algorithm is an iterative, gradient search, learning algorithm, which adjusts each weight in a multi layer net so as to reduce the error in the outputs. It works by propagating errors backward from the output layer. The global error in the performance with a given set of weights is the sum over all output units of the sum over all training cases of the squared distances between the actual and desired outputs of a unit. For a three layer ANN with five inputs, one hidden layer with four neurons and two output neurons, the error function is

$$E = \frac{1}{2} \sum_{k=1}^{2} \sum_{p=1}^{p} (y_{kp} - d_{kp})^2$$

where $y$ and $d$ are the desired and actual outputs of the $p$ training patterns, respectively.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Performance evaluation of DT of Fig. 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DT 1 – short circuit (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Success rate</td>
<td>95.319</td>
</tr>
<tr>
<td>False alarm</td>
<td>3.329</td>
</tr>
<tr>
<td>Missed alarm</td>
<td>5.871</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Performance evaluation of DT of Fig. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DT 2 – machine trip (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Success rate</td>
<td>96.527</td>
</tr>
<tr>
<td>False alarm</td>
<td>4.684</td>
</tr>
<tr>
<td>Missed alarm</td>
<td>3.141</td>
</tr>
</tbody>
</table>
Individual weight adjustments between the hidden layer and the output layer for pattern $p$ are computed by

$$\Delta w_{jk} = -l_r \frac{\partial E_p}{\partial w_{jk}}$$ (3)

where $j = 1, 2, 3, 4$ and $l_r$ is a constant learning rate. Through some derivations, the recursive formula for adjustment of weights with momentum term is obtained for the weights between hidden and output layers as it is given below:

$$w_{jk}(n + 1) = w_{jk}(n) + l_r \Delta w_{jk} + m \Delta w_{jk}(n - 1)$$ (4)

In Eq. (4), $n$ is the iteration number and $m$ is the momentum constant. Same weight adjustment formulas can be obtained between the input and output layers by changing the indices from $k$ to $I$.

A number of test patterns (1987 Ops) different from the training patterns were given to the developed ANNs in order to evaluate the accuracy of the results. The absolute error for each Ops of the first disturbance does not exceed the 0.18 Hz, as it is shown in Fig. 5.

The quality of the results needs to be evaluated by quantifying the mean relative error, the mean absolute error and the mean square error relatively to the target output values $y$, in this case the minimum frequency deviation of the power system—$f_{\text{min}}$. The performance evaluations for both disturbances are shown in Tables 3 and 4.

The results have shown that ANN based methods can provide a very good security assessment, giving an accurate estimation of minimum frequency values in case of some common disturbances [11].

5. Entropy networks

One popular type of neural networks are the multi-layer perceptrons (MLPs), which are feed-forward neural networks with one input layer, one or more hidden layers and one output layer [3]. In the case of two hidden layers, the first hidden layer is the partitioning layer or test layer (TL), which divides the entire feature space into several regions. The second hidden layer is the AND layer (AL), which performs adding of the partitioned regions. The output layer is the OR layer (OL), which combines the previous layer results to produce disjoint region of arbitrary shape.

It can be seen that decision trees (DTs) and multi layer perceptrons (MLPs) are equivalent in terms of input-output relations, as it is described in [7]. Thus, a

| Table 3 |
| Performance evaluation in first case |
| ANN 1 – Dist. (machine-loss) (%) |
| Success rate | 98.25 |
| Mean relative error | 0.036 |
| Mean absolute error | 1.749 |
| Mean square error | 4.114 |

| Table 4 |
| Performance evaluation in second case |
| ANN 2—Dist. (short circuit) (%) |
| Success rate | 97.38 |
| Mean relative error | 0.054 |
| Mean absolute error | 2.616 |
| Mean square error | 5.537 |
DT can be reformulated into a neural network, called entropy network, following the next steps for the formation of:

1. The input layer (IL): it contains one neuron per attribute selected and tested by the corresponding DT. In our application, there are seven system attributes selected in first DT and four attributes in the second.
2. The first hidden layer (TL): it contains one neuron per DT test node. Each TL neuron is linked to the IL neuron corresponding to the tested attribute. The test nodes for the first DT are 11, and for the second DT 8.
3. The second hidden layer (AL): it contains one neuron per DT terminal node. Each AL neuron is connected to the TL neurons corresponding to the test nodes located on the path from the top node towards the terminal node. There are 12 terminal nodes for the first DT and 9 for the second, respectively.
4. The output layer (OL): it contains one neuron per DT class, connected to the AL neurons corresponding to the DT terminal nodes, where the class is the majority class. Its activation is high, if at least one of these is active. Thus, all weights arriving at an OL neuron are equal to 1, and its threshold is equal to its number of inputs minus one.

The entropy networks (ENs) may be used to approximate a continuous security margin, rather than to merely classify. In this case, the above described output layer would be replaced by a single output neuron, fully connected to all neurons of the AL, while the weights should be recalculated through retraining. The developed ENs are illustrated in Figs. 6 and 7. In these cases, the neurons are represented by a circle, within which the corresponding DT node number and type are indicated.

Once the network structure is defined, these are trained by adapting their weights and thresholds to the input/output pairs observed in the LS. Each layer has a weight matrix $W$, a bias vector $b$ and an output vector that is input for the next layer. The two ENs have seven and four inputs ($I_i$), respectively, and two outputs ($O_i$). The total number of weights is 76, ([11] + [53] + [12]) for the first network and 60, ([8] + [33] + [9]) for the second. Back propagation algorithm was used for learning procedure, while the networks were trained with constant learning rate of 0.1, momentum of 0.06 and maximum error tolerance of 0.001.

As in the case of the DTs, the classification success is evaluated by quantifying the success rate, the false alarms and the missed alarm. As in case of DTs evaluation, a number of test patterns (1987 Ops)

<table>
<thead>
<tr>
<th>EN 1 – short circuit (%)</th>
<th>Success rate</th>
<th>False alarm</th>
<th>Missed alarm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>98.01</td>
<td>1.73</td>
<td>2.14</td>
</tr>
</tbody>
</table>
different from the training patterns were given to the developed ENs, in order to evaluate the accuracy of the results. The results are shown in Tables 5 and 6. These rates are compared to the corresponding rates of the DTs performance evaluation.

It is shown that the performance of the ENs is slightly improved, compared with the DTs performance. This means that the application of the ENs increases the reliability of the approach, while fast computational time, which is very important for on-line applications, is maintained.

### Table 6
Performance evaluation in second EDT

<table>
<thead>
<tr>
<th>EN 2 – machine trip (%)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Success rate</td>
<td>98.43</td>
<td></td>
</tr>
<tr>
<td>False alarm</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td>Missed alarm</td>
<td>2.61</td>
<td></td>
</tr>
</tbody>
</table>

6. **On-line program implementation**

All the previous developed algorithms has been embedded in a real-time performance energy management program, named CARE. The CARE software comprises various modules for short-term load and wind power forecasting, unit commitment, economic dispatch and on-line security assessment oriented to the needs of isolated power systems with increased renewable power penetration [8]. The security evaluation structures were integrated as modules, activated “on call” by the operators. In the execution cycle, security assessment follows the unit commitment and dispatch modules, leaving to the operator the decision to activate the module for validation of the proposed dispatch scenario. In Fig. 8 the dynamic security assessment module, as implemented in the EMS of Crete, is shown. Results for 24 h ahead are

![Fig. 8. Man–machine interface of the dynamic security assessment module.](image-url)
displayed in the form of lines that represent the frequency that is expected in case of the considered disturbances under the predicted load demand and wind power production, also displayed in the same diagram. In this case the maximum frequency deviation appears within the specified security limits for both disturbances.

Additionally, two more on-call screens of the program are presented in the following Fig. 9. First, the results of the developed short-term operation-planning module (unit commitment) for a given forecasting load demand are plotted, taking into account the technical restrictions and the economic characteristics of all the available power units. In the next window of the same figure, the corresponding dynamic security level of the system are estimated by the available DSA algorithms, in order to assess the dynamic behavior in case of same serious disturbances.

7. Conclusions

This paper describes the application of automatic learning techniques to the evaluation of the dynamic security of isolated power systems. These techniques have been integrated in the dynamic security assessment module of the advanced control system of the island of Crete, helping to identify the operating conditions and parameters that lead to a less robust operation of the system. The results are considered highly satisfactory, both in accuracy that increases the reliability of the method and in computational time, which is a necessity for real time applications.

Specifically, three machine-learning approaches have been applied and tested for on-line dynamic security assessment of the studied power system. DTs are first developed and applied in order to identify the most important system parameters for the given
problem. It has been shown that DTs are indeed an effective way to extract interpretable rules from very large bodies of simulated examples [12]. These operating rules were taken in account for the determination of the input variables for the ANNs, which provided a slightly better performance at similarly fast computational time, however, at the expense of results interpretability. Finally, entropy networks, a hybrid neural network approach was applied aiming at combining the attractive features of two techniques, namely the simplicity and transparency of decision trees and the information accuracy of multi layer perceptrons. In addition, the comparison between DTs and ENs shows that both techniques are able to provide an operator with an on-line classification of the system security, but DTs exhibit high interpretability and fast computational time, while ENs maintain these qualities and provide improved performance.

Acknowledgment

The authors wish to thank the European Commission DGXII for financing the R&D project: “MORE CARE: More Advanced Control for Secure Operation of Isolated Power Systems with Increased Renewable Energy Penetration and Storage”, contract ERK5-CT-1999-00019.

References


E.S. Karapidakis was born in Crete island, on August 8, 1972. He received the diploma degree in electrical and computer engineering in 1997 and his PhD degree in 2003 from the National Technical University of Athens (NTUA). Since then, he is a project evaluation and implementation manager of company for support and development of Cretan Enterprises, while his research concerns application of artificial intelligence techniques to power systems energy management.