Machine Learning Approaches to Power System Security Assessment

Louis WEHENKEL
Research Associate F.N.R.S. - Dept. of Electrical Engineering -Institut Montefiore
University of Liège - Sart Tilman, B28 - B 4000 Liège BELGIUM
lwh@montefiore.ulg.ac.be

Abstract

This paper describes ongoing research and development of machine learning and other complementary automatic learning techniques in a framework adapted to the specific needs of power system security assessment. In the proposed approach, random sampling techniques are considered to screen all relevant power system operating situations, while existing numerical simulation tools are exploited to derive detailed security information. The heart of the framework is provided by machine learning methods used to extract and synthesize security knowledge reformulated in a suitable way for decision making. This consists of transforming the data base of case by case numerical simulations into a power system security knowledge base. The main expected fallouts with respect to existing security assessment methods are computational efficiency, better physical insight into non-linear problems, and management of uncertainties. The paper discusses also the complementary roles of various automatic learning methods in this framework, such as decision tree induction, multilayer perceptrons and nearest neighbor classifiers. Illustrations are taken from two different real large scale power system security problems : transient stability assessment of the Hydro-Québec system and voltage security assessment of the system of Electricité de France.

Keywords: Electric power systems; security assessment; decision tree induction; neural networks; nearest neighbor.

1 Introduction

Security assessment is a major topic in planning and operation of electric power systems. It consists of evaluating the ability of the power system to face various contingencies and of proposing appropriate remedial actions able to counter its main weaknesses, whenever deemed necessary. Contingencies may be external or internal events (e.g. faults subsequent to lightning vs operator initiated switching sequences) and may consist of small/slow or large/fast disturbances (e.g. random behavior of the demand pattern vs generator or line tripping).

The effect of a contingency on a power system in a given state is usually assessed by numerical (e.g. time-domain) simulation of the corresponding scenario. However, the nonlinear nature of the physical phenomena and the growing complexity of real-life power systems make security assessment a difficult task. For example, the everyday monitoring of a power system calls for fast analysis, sensitivity analysis (which are the salient parameters driving the phenomena, and to which extent?), suggestions to control. On the other hand, increasing economic and environmental pressure make the conflicting aspects of security and economy even more challenging. Overall, the need for methods different from the standard time domain simulation is increasingly felt.

This paper describes ongoing research and development of such methods, using machine learning (and other automatic learning) techniques in a framework adapted to the specific needs of power system security assessment. In the proposed approach, schematically sketched in Fig. 1, random sampling techniques are considered to screen all relevant situations in a given context, while existing numerical simulation tools are exploited - if necessary in parallel - to derive detailed

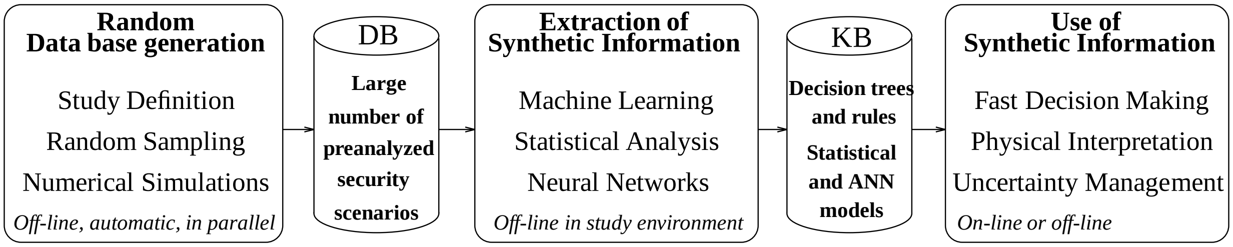


Figure 1: Machine learning framework for security assessment

security information. The heart of the framework is provided by machine learning methods used to extract and synthesize relevant information and to reformulate it in a suitable way for decision making. This consists of transforming the data base (DB) of case by case numerical simulations into a power system security *knowledge base* (KB). As illustrated in Fig. 1, a large variety of automatic learning methods may be used here in a toolbox fashion, according to the type of information they may exploit and/or produce. The final step consists of using the extracted synthetic information (decision trees, rules, statistical or neural network approximators) either in real-time, for fast and effective decision making, or in the off-line study environment, so as to gain new physical insight and to derive better system and/or operation planning strategies.

How will this automatic learning based framework complement classical system theory oriented methods (relying on analytical power system models, such as numerical simulation) for security assessment? In practice, there are three dimensions along which we expect important fallouts.

First of all *computational efficiency*. By using synthetic information extracted by automatic learning, instead of analytical methods, much higher speed may be reached for real-time decision making. Further, in terms of data requirements, whereas analytical methods require a full description of the system model, the approximate models constructed via automatic learning may be tailored in order to exploit only the significant input parameters. Computational efficiency was actually the motivation of Dy Liacco, when he first envisioned in the late sixties the use of automatic learning (at that time, statistical pattern recognition) for real-time security assessment [1]. Even today, and in spite of the very significant increase in computing powers in the last twenty-five years, this remains a strong motivation.

But the synthetic information extracted by automatic learning methods, may itself be complementary to and generally more powerful than that provided in a case by case fashion by existing analytical methods. In particular, much more attention is paid nowadays to *interpretability* and management of *uncertainties*, the two other important fallouts expected from automatic learning methods.

As concerns *interpretability*, the use of automatic learning to provide physical insight into the nonlinear system behavior was first proposed by Pao et al in the mid-eighties [2]. In the meanwhile, it has been demonstrated that machine learning is indeed an efficient and effective way to generate reliable and interpretable security rules from very large bodies of simulated examples [3, 4], even for as complex systems as are real large-scale power systems. The extracted rules are found to express explicitly problem specific properties, similarly to human expertise, and hence may be easily appraised, criticized and eventually adopted by engineers in charge of security studies. This means that the above framework should also be viewed as an approach to the maintenance and enhancement of utility expertise. The flexibility of the machine learning framework allows one to tailor the resulting information to analysis, sensitivity analysis and control applications.

As concerns management of *uncertainties*, the need to devise a rational way to take decisions whenever there are major uncertainties about the power system state becomes more and more apparent. Today, for example, it is well known that operators are often sorely missing guidance in the context of unusual system states reached after major disturbances, where reliable real-time information is generally lacking. Tomorrow, technological and economic changes will probably lead to a higher and physically more irrational distribution of decision making and thus to more uncertainties in routine operation and planning activities. Indeed, on the one hand, new devices (e.g. flexible alternating current transmission systems (FACTS)) may cause stronger interactions among remote components of very large interconnections. On the other hand, increased competition among economic actors may further reduce their willingness to share information on their respective subsystems, in spite of the stronger physical interactions. Under such circumstances, approaches able to manage uncertainties, such as the above framework based on automatic learning, will be urgently needed.

Nonetheless and despite repetitive attempts, there are still no large-scale industrial applications of the machine learning framework to power system security assessment. This is mainly due to the fact that until recently, the existing automatic learning methods were not powerful enough while the amount of possible security studies was limited by available simulation hardware and software.

Today, however, all the required conditions are met. Present day computer networks together with fast simulation tools allow the generation of large amounts of detailed studies. At the same time much progress has recently been achieved in automatic learning methods and their application to large-scale power systems was shown to be feasible. Hence, automatic information synthesis tools to assist engineers to compare and interpret the numerous elementary results, and extract and appraise useful synthetic information are at the same time strongly needed and technically feasible.

Therefore, while we expect additional progress in learning methods and application methodologies, we foresee that some important electric power companies e.g. in North America or Europe will soon start using this approach more or less routinely for security studies.

2 Aspects of power system security problems

In this section we provide a guided tour of power system security for the unfamiliar reader. We will first analyze the different types of physical problems, then consider the practical application environments where security is treated, and finally mention briefly the main classes of existing analytical tools for security assessment. In our discussion, we will focus on security problems involving *large* disturbances corresponding to nonlinear system behavior. Although such disturbances are generally very unlikely to happen, their potential consequences can be extremely important and may lead to complete system blackouts, freezing the economic activity of a whole country for many hours.

2.1 Classification of operating states

The different operating modes of a power system were defined by Dy Liacco [1]. Figure 2 shows a more detailed description of the “Dy Liacco state diagram”.

Preventive security assessment is concerned with the question whether a system in its normal state is able to withstand every plausible disturbance, and if not, preventive control would consist of moving this system state into a secure operating region. Since predicting future disturbances is difficult, preventive security assessment will essentially aim at balancing the reduction of the *probability* of losing integrity with the economic cost of operation.

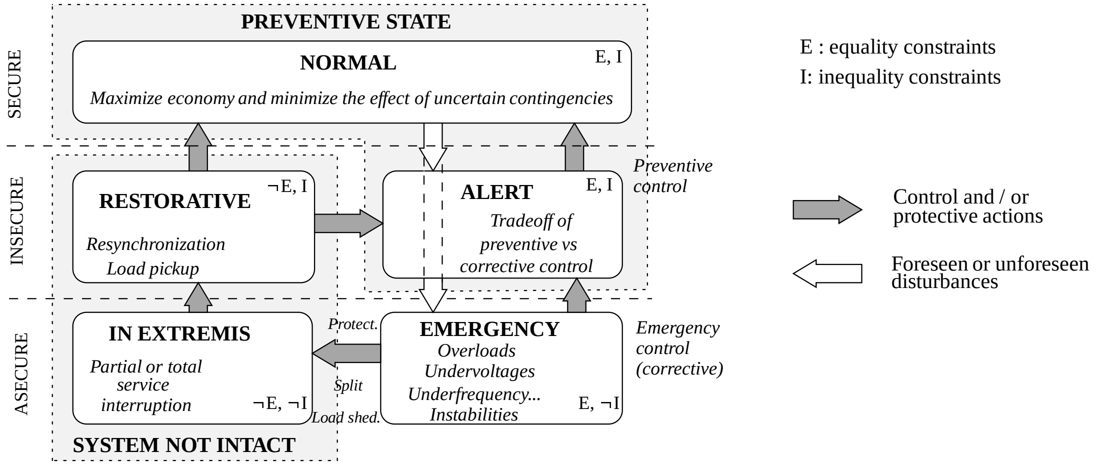


Figure 2: Operating states and transitions. Adapted from [5]

Emergency state detection aims at assessing whether the system is in the process of losing integrity, following an actual disturbance inception. This is a more deterministic evolution, where response time is critical while economic considerations become temporarily secondary. Emergency control aims at taking fast last resort actions, to avoid partial or complete service interruption.

When both preventive and emergency controls have failed to bring system parameters back within their inequality constraints, automatic local protective devices will act so as to preserve power system components operating under unacceptable conditions from undergoing irrevocable damages. This leads to further disturbances, which may result in system splitting and partial or complete blackouts.

Consequently, the system enters the restorative mode, where the task of the operator is to minimize the amount of un-delivered energy by resynchronizing lost generation as soon as possible and picking up the disconnected load, in order of priority.

We will confine ourselves to preventive and emergency aspects.

2.2 Physical classification of security problems

Various security problems are distinguished according to the time scales of the dynamics, the characteristic symptoms (low voltage, large angular deviations. . .), and the control means (reactive power, switching. . .) to alleviate problems.

Transient stability. Transient stability concerns the ability of the generators of a power system to recover synchronous operation following the electromechanical oscillations caused by a *large* disturbance. In this context, the dynamic performance is a matter of seconds and is mainly affected by switching operations and fast power controls (e.g. fast valving, high voltage direct current converters, FACTS) and voltage support by the automatic voltage regulators of synchronous generators and static var compensators (SVCs). To determine the *degree* of stability we may evaluate the critical clearing time of a fault, which is the maximum time duration it may take to clear the fault without the system losing its ability to maintain synchronism.

Voltage security. The fastest voltage instabilities are characterized by sudden voltage collapse phenomena which may develop at the same or even higher speeds than loss of synchronism. More classical is the *mid-term* voltage instability, which corresponds to a typical time frame of one to

Table 1: Security assessment environments. Adapted from [6]

Environment	Time scales	Typical problems	Operator	Expert
System planning	1 - 10 years	Generation Transmission Protection	No	Yes
Operation planning	1 week - 1 year	Maintenance Unit commitment Protection settings	No	Yes
On-line operation	1 hour - 1 day	Preventive mode Security assessment	Yes	Partly
Real-time* monitoring	sec. - min. - hour	Emergency control Protective actions	No**	No
Training	months - days	Improve operator skill	Yes	No

* Here we distinguish between *real-time*, which considers dynamic situations following a disturbance inception, from merely *on-line* which considers static pre-disturbance situations.

** except for static security corrective control

five minutes. In this case voltage collapse is mainly driven by automatic transformer on-load tap changers trying to restore voltage nearby the loads. There is a third, even slower time frame, corresponding to the so-called *long-term* voltage instability, which involves the gradual buildup in load demand. This interacts with classical static security and is well within the scope of operator intervention.

Although a voltage collapse may result in a wide spread degradation of the voltage profile and subsequent loss of synchronism, it is normally initiated by a *local* deficiency in reactive power reserves and/or a reduced reactive power transmission capability into a given load area. The distance to voltage insecurity may be evaluated by a load power margin which is the maximum additional amount of power which may be transferred safely from the generation to a given load area.

Static security. It concerns essentially thermal overload problems of generation transmission system components, where phenomena span over significantly longer periods of time. For example, line overloads may be tolerated during 30 to 60 minutes under favorable weather conditions.

2.3 Practical application domains

Table 1 shows the practical study contexts or environments which may be distinguished in security assessment applications. The first column identifies the study context; the second specifies how long in advance (with respect to real-time) studies may be carried out; the third column indicates the type of subproblems that are generally considered in a given environment; the last two columns indicate respectively if an operator is involved in the decision making procedure and if an expert in the field of power system security is available.

In the first three contexts one currently relies mostly on the intervention of human experts exploiting the numerical simulation tools. In real-time monitoring and emergency control, the reduced time available calls for more automatic procedures.

System planning. Multitudinous system configurations must be screened for several load patterns, and for each one a large number of contingencies. An order of magnitude of 100,000 different scenarios per study would be realistic for a medium sized system. While enough time may be available to carry out so many security simulations, there is still room for improved data analysis methods to exploit their results more effectively for the identification of structural system weaknesses and to provide guidelines to improve reliability.

Operation planning. As suggested in Table 1, operation planning concerns a broad range of problems, including maintenance scheduling (one year to one month ahead), design of operating strategies for usual and abnormal situations, and setting of protection delays and thresholds. The number of combinations of situations which must be considered for maintenance scheduling is also generally very large, and automatic learning approaches would equally be useful to make better use of the available information and to exploit the system more economically.

Similarly, for the closer to real-time determination of operating security criteria, machine learning is particularly well adapted. It would allow engineers to screen more systematically representative samples of situations, in order to identify critical operating parameters and determine their security limit tables needed for on-line operation. This would actually consist of automating and enhancing such manual approaches presently in use at many utilities.

On-line operation. In the context of this framework, it would consist of exploiting on-line the security knowledge bases set up off-line, e.g. in operation planning. Appropriate strategies are required in order to update this information when major changes happen in the system. For example, several weeks ahead routine security criteria could be designed for a forecast range of topologies, load levels and generation schedules, while, closer to real-time, maybe a day or some hours ahead, these criteria might then be refreshed to handle previously unexpected situations. In order to be compatible with the way operators usually appraise their system, it is particularly important for the synthetic information extracted by automatic learning to be as simple as possible to interpret.

Real-time monitoring. Here, the purpose is to design criteria to trigger more or less automatically emergency control actions, so as to prevent a disturbed system state to evolve towards blackout. An important aspect is the use of appropriate models[1] to reflect the *disturbed* power system behavior, when designing the security criteria. Furthermore, the use of readily available system *measurements* as inputs to the derived emergency control rules is often an operational constraint in addition to minimal data requirements and ultra high speed.

Training. During operator training, the security criteria derived in either of the preceding contexts might be usefully exploited as guidelines, provided that they are presented in an intelligible way. In addition, these models might be used internally in a training simulator software, in order to set up particular scenarios presenting particular insecurity modes.

2.4 Analytical tools

A rather large set of numerical methods are available for security assessment in the different time frames mentioned. We call them *analytical* tools since they exploit analytical power system models in contrast to the *synthetic* ones extracted by automatic learning techniques. Some of them are based on general purpose power system dynamic simulation packages and have a very broad scope. Others are based on simplified models or approaches representing only those features which are relevant for the particular study. The latter methods may be significantly more efficient, although at the expense of being restricted to some particular physical phenomena and/or some particular

[1] Depending upon the context, we use the term “model” either to denote the physical power system model or the synthetic information extracted by automatic learning.

(types of) power systems. We briefly discuss them since they provide the raw input data exploited by the automatic learning methods in order to synthesize the high level security information.

2.4.1 Transient stability

There are two main classes of analytical tools for transient stability assessment : time-domain (or step-by-step) simulation approach and direct methods, based on the second Lyapunov method.

Time-domain simulation. The general power system dynamic model is composed of mixed algebraic and differential equations strongly nonlinear, involving typically a few thousand discrete or continuous time state variables. To assess transient stability, the time-domain approach consists of simulating the during and post-fault behavior of the system for a given disturbance, and observing its electromechanical angular and voltage swings during a few seconds. Practical criteria vary from one utility to another, but an unacceptable performance would generally imply too large or undamped angular deviations (e.g. pole slips) or excessively large variations of voltage or frequency. To obtain stability margins, repetitive simulations must be carried out for various pre-fault operating states or for various assumptions concerning the delays of protection devices. While this approach is still considered as very CPU intensive, we observe that within the last three years the time required for a typical power system simulation with high order models has shrunk from one hour to some minutes.

Direct Lyapunov type methods. They aim at identifying when the system leaves its stability domain without further integration of the system trajectory. By avoiding the simulation of the post-fault trajectory, they reduce the simulated time period to a fraction of a second instead of the several seconds of time-domain methods. Some of them are thus able to provide a rich stability assessment (margins, sensitivities, mode of instability) within a fraction of the time required for a single time-domain simulation. A major drawback is their difficulty to exploit accurately models of generators and control loops as well as nonlinear or dynamic loads. However, since the first multimachine direct methods developed in the late sixties much progress has been achieved in incorporating more realistic models.

2.4.2 Voltage stability and security

Tools for voltage security assessment range from static load-flow calculations to full short-term / mid-term time domain simulations. It is worth mentioning that due to the rather recent emergence of voltage security problems, modeling practices have not yet reached maturity comparable to those used in transient stability studies. In particular, one intrinsic difficulty of analyzing voltage collapse phenomena is the very strong dependence on load behavior, for which good models are generally missing.

Short-term / mid-term dynamic simulations. Since voltage collapse phenomena may involve time constants ranging from a fraction of a second to a few minutes, a variable step-size numerical integration method with stiff system simulation capability is preferable for the sake of efficiency and accuracy, in contrast to transient stability where fixed step-size methods have been widely used.

Simplified simulations. Since many voltage security problems are essentially driven by automatic on-load tap changer mechanisms, it is possible to neglect sometimes the faster interactions among load and generation dynamics. The differential equations corresponding to the faster phenomena are then replaced by equilibrium equations and only the slower dynamics are modeled. With the intrinsic limitation of neglecting problems caused by the fast dynamics, this kind of approach allows drastic reduction in computing times.

Post-contingency load-flow. A further simplification consists of neglecting totally the dynamics, and using only purely static post-contingency load-flow calculations. Typically, this allows one to compute maximal loading limits, based on successive computations or even on direct optimization.

3 Aspects of automatic learning

In this section we introduce classes of potentially useful automatic learning methods for the synthesis of security assessment information. We first give a definition of the generic *supervised* learning problem and introduce three important classes of algorithms for this problem, and finish with some comments on the use of *unsupervised* learning methods.

3.1 Supervised learning problem

The generic problem of supervised learning from examples can be formulated as follows :

> *Given a learning set of examples of associated input/output pairs, derive a general model for the underlying input/output relationship, which may be used to explain the observed pairs and/or predict output values for any new unseen input.*

In the context of security assessment, an example corresponds to a given operating situation. The input attributes would be (hopefully) relevant parameters describing its electrical state and topology and the output could be information concerning its security, in the form of either a discrete classification (e.g. secure / marginal / insecure) or a numerical value derived from security margins.

In general, the solution of this overall learning problem is decomposed into several subtasks.

Representation consists of (i) choosing appropriate input attributes to represent the power system state, (ii) defining the output security information, and (iii) choosing a class of models suitable to represent input/output relations.

The *representation problem* is left to the engineer. A compromise has to be found between the use of very elementary standard operating parameters and more or less sophisticated compound features. Below we discuss how unsupervised learning techniques may help to choose appropriate input attributes.

Feature selection aims at reducing the dimensionality of the input space by dismissing attributes which do not carry useful information to predict the considered security information. This allows one to exploit the more or less local nature of many security problems.

Model selection (or learning per se) will typically identify in the predefined class of models the one which best fits the learning states. This generally requires choice of model structure and parameters, using an appropriate search technique.

The distinction between *feature selection* and *model selection* is somewhat arbitrary, and some of the methods actually solve these two problems simultaneously rather than successively.

Interpretation and validation are very important in order to understand the physical meaning of the synthesized model and to determine its range of validity. It consists of testing the model on a set of unseen test examples and comparing its information with prior expertise about the security problem.

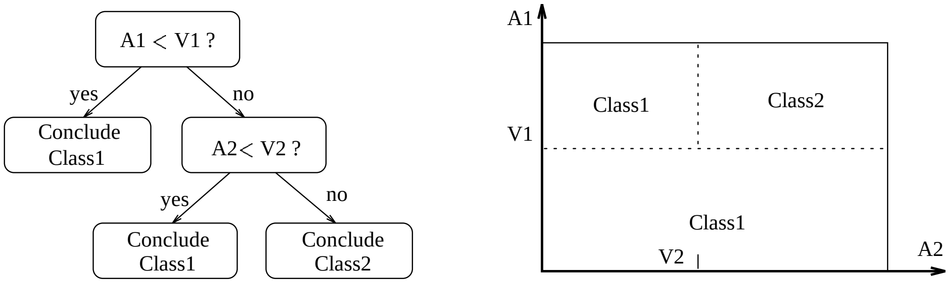


Figure 3: Hypothetical decision tree and its corresponding input space decomposition

From the *interpretation and validation* point of view, some supervised learning methods provide rather black-box information, difficult to interpret, while some others provide explicit and very transparent models, easy to compare with prior knowledge.

Model use consists of applying the model to predict security of new situations on the basis of the values assumed by the input parameters, and if necessary to "invert" the model in order to provide information on how to modify input parameters so as to achieve a security enhancement goal.

As far as the *use* of the model for fast decision making is concerned, we notice that there are speed variations of several orders of magnitude between various techniques, but most of the methods are sufficiently fast in the context of control center oriented power system security analysis.

3.2 Supervised learning methods [7]

In what follows, we consider only non-parametric automatic learning methods. Parametric methods may be useful in some particular circumstances, but are not powerful enough to treat the wide variety of practical security problems. We will discuss three classes of methods providing three *complementary* types of information. Although we have selected them from three different paradigms (machine learning, neural nets, pattern recognition) we insist on the type of information provided rather than on the paradigm itself.

3.2.1 Symbolic knowledge via machine learning

Machine learning is the subfield of artificial intelligence concerned with the design of automatic procedures able to learn from examples. *Concept learning from examples* denotes the process of deriving a logical description of the necessary and sufficient conditions corresponding to a class of objects, i.e. a *rule* in some given representation language. A major concern is to find out adequate compromises between rule *complexity* and data fit, so as to avoid over-fitting and to privilege interpretability.

Top down induction of decision trees (TDIDT) is one of the most successful classes of such methods which was popularized by Quinlan [8]. Figure 3 shows a hypothetical binary decision tree (DT) : to infer the output information corresponding to given input attribute values, one traverses the tree, starting at the top-node, and applying sequentially the dichotomous tests encountered to select the appropriate successor. When a terminal node is reached, the output information stored there is retrieved.

As suggested by the acronym, TDIDT approaches the decision tree learning in a divide and conquer fashion, whereby a decision tree is progressively built up, starting with the top-node and ending up with the terminal nodes. At each step, a tip-node of the growing tree is considered and the algorithm decides whether it will be a terminal node or should be further developed. To develop a node, an appropriate attribute is first identified, together with a dichotomy on its values. The subset of its learning examples corresponding to the node is then split according to this dichotomy into two subsets corresponding to the successors of the current node. The terminal nodes are “decorated” with appropriate information on the output values derived from their learning examples, e.g. the majority class label or probabilities, or expected value and standard deviation of numerical output information.

The right part of Fig. 3 shows how the decision tree in its left decomposes its input space into *non-overlapping* subregions. The number of such regions should ideally be as small as possible and at the same time the states contained by each region should belong to a same class. Thus, to build good decision trees, an algorithm must rely on appropriate *optimal splitting* and *stop splitting* rules. Optimal splitting has to do with selecting a dichotomy at a test node so as to provide a maximum amount of information on the output value (i.e. separate states of different classes) whereas stop splitting has to identify situations where further splitting would either be useless or lead to performance degradation, due to over-fitting.

Decision trees have been quite extensively studied in the context of various security assessment problems [6]. A main asset lies in the explicit and logical representation of the induced classification rules and the resulting unique explanatory capability. In particular, the method provides systematic correlation analyses among different attributes and identifies the most discriminating attributes at each tree node. From the computational viewpoint it is efficient at the learning stage as well as at the prediction stage.

There are two generalizations of decision trees of interest in the context of security assessment. First, *regression* trees which infer information about a numerical output variable; they are illustrated below. Second, *fuzzy* trees which use fuzzy logic instead of standard logic to represent output information in a smooth fashion. Both approaches allow us to infer information about security margins, similarly to the techniques discussed below. Fuzzy trees have not yet reached the maturity of crisp classification or regression trees, but they seem particularly well suited to our types of problems. Indeed, they appear to be more robust with respect to noise than classical machine learning methods and are able to combine smooth input/output approximation capabilities of neural networks with interpretability features of symbolic machine learning [9].

3.2.2 Smooth nonlinear approximations via artificial neural networks

The field of artificial neural networks has grown since the early work on perceptrons to an important and productive research field. We restrict ourselves to multilayer perceptrons; for further information, a widely recommended theoretical introduction to neural networks is given in [10].

The single-layer perceptron, is basically a simple linear threshold unit together with an error correcting learning algorithm. It is able to represent a linear boundary in its input space. Its limited representation capabilities have motivated the consideration of more complex models composed of multiple interconnected layers of perceptrons, MLPs for short. Figure 4 illustrates the classical feed-forward MLP. The first or *input* layer corresponds to the attribute values, and the last or *output* layer to the desired security classification or margin information. Intermediate layers enable the network to approximate arbitrarily complex input/output mappings, provided that its topology and its weights are chosen appropriately.

The discovery of the back-propagation algorithm has been central to the success of MLPs. It allows one to compute efficiently and locally the gradient of the output error of the network with respect to its weights and thresholds. It may be exploited iteratively in order to adjust the weights so as to

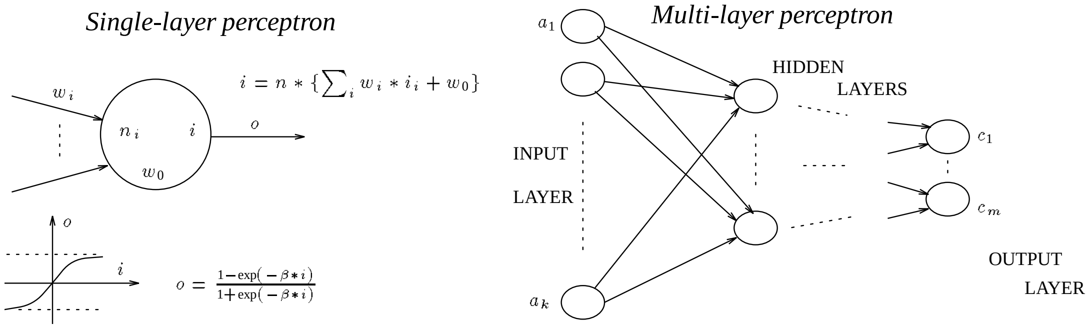


Figure 4: Feed forward multilayer perceptron

reduce the total mean square output error for learning examples. In recent years, much progress has been made in improving efficiency of optimization techniques for the learning procedures of MLPs, but the MLPs are still very slow at the learning stage, which may prevent extensive experimentations for data base sizes typical of security assessment of realistic power systems.

Similarly to decision trees, an interesting property of MLPs is their ability to achieve feature extraction and learning in a single step : the weights connecting the input layer with the first hidden layer may be interpreted as projecting the input vector in some particular directions, realizing a linear transformation of the input space, which is used in subsequent layers to approximate outputs. However, one of the difficulties with MLPs comes from the very high number of weights and thresholds related in a nonlinear fashion, which makes it almost impossible to give any insight into the relationship learned. All in all, one can say that MLPs offer a flexible, easy to apply, but essentially black-box type of approach to function approximation.

It should be observed that a bunch of similar methods exist nowadays, such as radial basis functions and projection pursuit regression techniques. They offer the possibility of translating the case by case information provided in the learning sets into an approximate but closed form numerical model. The latter one may be used for fast assessment of unseen situations and direct computation of sensitivities.

3.2.3 Memory based reasoning via statistical pattern recognition [11]

The previous two approaches essentially compress detailed information about individual simulation results into general, more or less global security characterizations.

Additional information may however be provided in a case by case fashion, by matching an unseen (e.g. real-time) situation with similar situations found in the data base. This may be achieved by defining generalized distances so as to evaluate similarities among power system situations, together with appropriate fast data base search algorithms.

A well known such technique is the "K nearest neighbors" ($K - NN$) method able to complete decision trees and multilayer perceptrons. It consists of classifying a state into the majority class among its K nearest neighbors in the learning set. The main characteristics of this method are high simplicity but sensitivity to the type of distances used. In particular, to be practical, ad hoc algorithms must be developed to choose the distances on the basis of the learning set. While in the past this method was generally exploiting a small number of sophisticated ad hoc input features manually selected on the basis of engineering judgment, nowadays the emphasis is more on the research of automatic distance design methods exploiting the learning states.

3.3 Clustering and unsupervised learning

In contrast to supervised learning, where the objective is clearly defined in terms of modeling the underlying correlations between some input variables and some particular output variables, unsupervised learning methods are not oriented towards a particular prediction task. Rather, they try to identify existing underlying relationships among a set of objects characterized by a set of variables or among a set of variables used to characterize a set of objects.

Thus, one of the purposes of clustering is to identify homogeneous groups of similar objects, in order to represent a large set of objects by a small number of representative *prototypes*. Graphical, two-dimensional scatter plots may be used as a tool in order to analyze the data and identify clusters. Another application of the same techniques is to identify similarities (and redundancies) among the different attributes used to characterize objects. In the context of power system security both applications may be useful as complementary data analysis and preprocessing tools.

Unsupervised learning algorithms have been proposed under the three umbrellas given above to classify classification methods, termed *cluster analysis* in the statistics literature, *conceptual clustering* in the machine learning community, and *self-organizing maps or vector quantization* in the neural net community [12].

4 Application of automatic learning to power system security

Below we will first describe a hypothetical application of the automatic learning based framework to a hypothetical security problem. Then we will provide a short overview of some real-life applications to large-scale security problems.

4.1 A hypothetical illustration of the framework

4.1.1 A security problem

Let us imagine that our hypothetical power system is voltage security limited in some reactive power weak area, and let us suppose this security problem was discovered in a preliminary screening security study, where also the possibly constraining disturbances were identified.

Then, a practical problem would be the characterization of security regions with respect to these disturbances, so as to provide operators with preventive security assessment criteria and effective preventive control means to alleviate potential insecurities, such as optimal rescheduling of available reactive power resources.

Another, different problem would be the design of emergency state indicators to be applied in case of a disturbance, ideally highly anticipative and reliable at the same time while providing information on appropriate emergency control means, such as on-load tap changer blocking and load shedding.

4.1.2 How could we generate a data base ?

In order to provide a representative sample of voltage security scenarios for the above problems, we would first ask for the advice of planning and operation planning engineers and operators of that system, so as to gather information about known system weaknesses and operating practices.

From this information, data base building software would then be designed in order to generate randomized samples representative of normal operating conditions, including also a sufficient number of unusual situations, deemed relevant for security characterization. In particular, with respect to real-life operating statistics, this sample would typically be biased towards the insecure regions of the state space.

According to that sampling procedure, an initial data base would be generated, typically comprising several thousand states and the security of each state would be pre-analyzed with respect to the studied disturbances. For example, post-contingency load power margins could be computed for real large-scale power system models on existing computer networks within some hours of response time, by using an efficient simulation software and exploiting trivial parallelism. In addition to this information, appropriate preventive or emergency control information could be pre-determined for the insecure states and secure economic generation dispatch for the secure ones.

Further, a certain number of attributes would be computed, which would be proposed as input variables to formulate security criteria. In the preventive mode security assessment problem, these attributes would typically be *contingency-independent* pre-fault operating parameters, such as voltages, reactive power generation and compensation reserves, power flows, topology indicators. For the emergency state detection problem, we would rather use raw system measurements (e.g. voltage magnitudes, power flows, transformer ratios, breaker status) of the intermediate *just after disturbance* state. In contrast to the preventive mode attributes, the emergency state attributes would depend on the disturbance and on the short-term system modeling, in addition to the pre-fault operating state.

When designing the data base generation software, care must be taken so as to appropriately take into account various kind of uncertainties. For example, random noise terms should be added to the attribute values so as to model measurement or state estimation errors and delays. Further, static and dynamic power system model parameters are often uncertain (load distribution and sensitivity to voltage, external systems, parameter variations with temperature . . .) and should thus be accordingly randomized.

4.1.3 Unsupervised learning for data pre-processing

In practical security problems, many different attributes often turn out to provide equivalent information, due to the very strong physical correlations among geographically close components of a power system. Thus, clustering methods may be used to define a small set of representative attributes from a larger number of elementary variables.

To fix ideas, let us consider the case of voltage magnitudes. Correlation coefficients among any pair of bus voltages may be easily computed on the basis of the data base statistical sample. They may then be used as similarity measures by a clustering algorithm searching for a reduced number of voltage “coherent” regions. For each region an equivalent (e.g. mean) voltage would be used as an attribute instead of individual bus voltages, and the computational burden of the subsequent supervised learning of security criteria would be reduced, while robustness and interpretability would be improved. For example, two-dimensional Kohonen feature maps may be exploited in order to visualize the relationships among voltage regions and compare them easily with the geographic location of busbars in the power system.

In addition to the above “feature extraction” application, clustering techniques have also been proposed in a more conventional way, to identify groups of similar power system operating states. One possible purpose is to partition a very large data base into smaller subsets for which the security assessment problem could be easier to solve. Another interesting application would be to “condense” the full data base into a reduced number of representative prototypes, thereby decreasing the number of required security simulations and shortening the associated computation delays.

4.1.4 Supervised learning of security criteria

Given a data base composed of examples, for which security margins have been determined for several contingencies and a number of candidate attributes have been computed, supervised learning would proceed so as to derive appropriate security criteria. First of all however, the data base would be partitioned into disjoint learning and test samples. The learning sample will be used to build the synthetic security criteria, whereas the test set will be used to assess their reliability by comparing the security information predicted by them and the "real" one determined by simulation. In addition to the unseen test states generated automatically together with the learning states, a test sample representative of *actual* operating statistics should be collected from historical on-line records.

What can decision trees do ? We need first to define security classes by appropriate thresholds on the security margin. Then, the decision tree building includes (i) the automatic identification of the subset of attributes among the candidate ones relevant for the prediction of the security class (say ten to twenty among one or two hundred), and (ii) the definition of appropriate threshold values for these attributes so as to provide an approximate model of the voltage security region of the studied area of the power system. In addition to a global DT covering all disturbances simultaneously, single-contingency DTs may also be constructed to provide more specific information and additional insight. Further, various DTs may be constructed for various security margin threshold values, so as to discriminate between marginally secure and very secure situations. Depending upon whether normal pre-disturbance or just after disturbance attribute values are used, the DTs can be used either in a preventive or in an emergency wise approach.

If there are too many non-detections of insecure states, the threshold value used to define the secure class in terms of the security margin may be increased before rebuilding a tree. If there are too many false alarms, additional candidate attributes or learning states should be used.

What can neural networks add ? In addition to the simplified view on security, provided by the DTs in terms of a discrete model relating a small number of security classes and thresholds on attribute values, one is generally interested in providing a continuous security margin, at least in the neighborhood of the threshold values used to define security classes.

As we have mentioned, one of the strong points of the MLP is its nonlinear modeling capability. On the other hand, the decision tree identifies the attributes in strong correlation with the security class. Thus, in a hybrid approach we may use the latter attributes as input variables to a MLP model, while using a normalized security margin as output information.

In practice it may be necessary to proceed by trial and error to determine an appropriate number of hidden neurons and topology for the MLP structure. Once its structure and weights have been adapted on the basis of the learning states, the MLP provides a closed-form and differentiable security approximator, which may be used for fast margin prediction for any seen or unseen state and as well to compute margin sensitivities to attribute values.

Practical experiments reported below with various security problems have shown that this leads to richer and more reliable security assessment information.

What do distance based methods offer ? With the previous two approaches, we have essentially compressed detailed information about individual simulation results into general, more or less global security characterizations. This allows us to provide the required physical understanding, thanks to the data analysis component of decision trees and attribute clustering techniques. In addition, the derived models may be used efficiently for on-line security analysis.

In this latter context, further information may be obtained via memory based reasoning exploiting appropriate distances to find the most similar pre-analyzed situations to the real-time state. Once identified, these may be used in multitudinous ways. For example, their distance to the current state would provide a measure of confidence of the security information provided by any model derived from the data base (DT and/or MLP). If the latter were too large, it would then be concluded that for the current state no reliable security information may be derived from the data base. If the nearest neighbors were on the contrary sufficiently close to the current state, then various kinds of detailed and specific security information may be extrapolated from these states to the current situation and shown to the operator, including detailed contingency analysis and preventive and/or emergency controls.

4.2 Overview of some real-life applications

Below we provide more specific information about feasibility studies of the automatic learning approach, made for some real practical power system security problems.

4.2.1 Transient stability

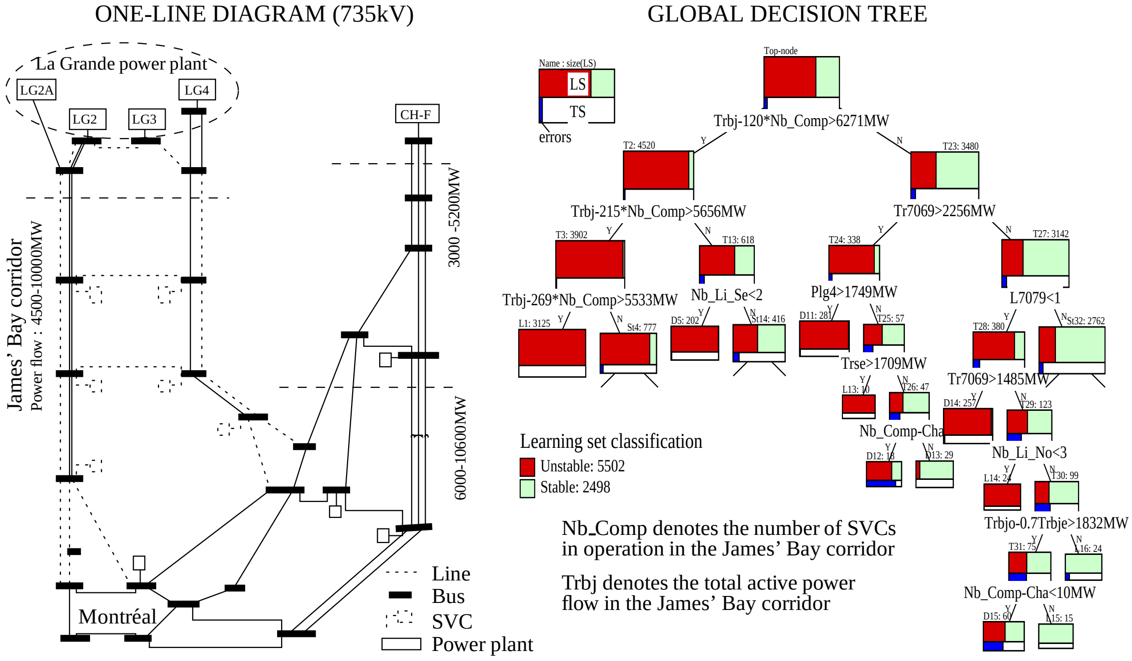
A first large-scale feasibility study was initiated in early 1990, for preventive transient stability assessment of an important generation plant within the large-scale EHV system of Electricité de France (EDF) [13].

A more recent study was carried out on the Hydro-Québec system which is illustrated in Fig. 5. Its normal operating condition is considered secure if it withstands any permanent single-phase to ground fault, followed by line tripping, fast re-closure and subsequent permanent tripping. It is notable that this system is mainly constrained by its transient stability limits, due to the very large power flows and long transmission distances.

More specifically, in our investigations we have considered only faults occurring within the James' Bay transmission corridor in the Western part of the system. With respect to such faults, the stability is mainly influenced by the power flows and topology within the same corridor. A set of transient stability limits have previously been developed, in a manual approach, where operation planning engineers have determined off-line, on the basis of carefully chosen simulation scenarios, a set of approximate limit tables relating the system topology and power flows to a Stable/Unstable classification. These limit tables have been implemented on the real-time computer of Hydro-Québec, via an ad hoc data base tool called LIMSEL, presently in use for operation. The purpose of our investigation was to evaluate the capability of the automatic learning approach to provide a more systematic and potentially more efficient methodology to derive these operating guidelines.

A data base, composed of 12,500 normal operating states was generated via random sampling and chained load flow computations; it comprises more than 300 different combinations of up to 6 line outages, and about 700 different combinations of reactive voltage support equipment in operation, and a wide variety of power flow distributions. The dashed lines in Fig. 5 show the variable topology part of the 735kV system. For each state, the corresponding classification Stable/Unstable was obtained from LIMSEL, running on the backup on-line computer, resulting in 3,939 stable states and 8,561 unstable ones, among which 393 are marginally and 8,168 fairly unstable.

To describe the operating states, and in order to characterize their stability, the following types of


ONE-LINE DIAGRAM (735kV)

GLOBAL DECISION TREE

Figure 5: Transient stability assessment of the Hydro-Québec system

candidate attributes were computed : active power flows through important lines and cut-sets in the James' Bay corridor; total active power generated in the 4 La Grande (LG) power plants and various combinations; number of SVCs in operation within the six substations in the James' Bay corridor; logical indicators (in/out) for important lines. This set, composed of 67 candidate attributes was determined with the help of an expert in charge of transient stability studies at Hydro-Québec. From previous studies it was already known that the total power flow through the corridor would be an important attribute, together with the topological information and the total number of SVCs.

The tree partially represented in the right hand part of Fig. 5 was built on the basis of the first 10,000 states [2] of the data base and 87 candidate attributes, including in addition to the above 67 ones four linear combination attributes and some other combined ones. Fig. 5 shows its most important parts nearby the top-node. The notation used for a typical node is also represented at the top left hand side of the tree : each node is represented by a box, the upper part of which corresponds to the proportions of stable and unstable *learning* states relative to this node. Test nodes are identified by the label "Ti" or "STi", the latter corresponding to subtrees which have not been drawn on the picture. Terminal nodes are identified by a label "Li" for leafs and "Di" for deadends. A leaf is a terminal node with a sufficiently class pure learning subset (in the algorithm used this is expressed in terms of an *entropy* measure) whereas a deadend is a node which corresponds to a subtree pruned to avoid over-fitting. In addition to the label indicating the type of a node, the number of learning states of the node is indicated next to it.

All in all, the tree comprises 57 test nodes and 58 terminal ones. It has identified among the candidate attributes the 24 most relevant ones. Among others, at several test nodes (including the top-node) the algorithm has selected a linear combination of the total power flow "Trbj" in the James' Bay corridor and the number of SVCs in operation "Nb_Comp" which thus confirms prior

[2] 8,000 are used to grow the DT and 2,000 to determine its optimal degree of pruning

knowledge. Thus, the threshold values of “Trbj” are function of “Nb_Comp”. For example, if “Nb_Comp”=12, the leftmost terminal node “L1” in Fig. 5 corresponds to a limit value of

$$\max\{6271 + 12 * 120, 5656 + 12 * 215, 5533 + 12 * 269\} = 8{,}761\text{MW},$$

above which a state is unconditionally declared unstable for at least one line-fault in the corridor.

To evaluate its generalization capability, the tree was tested on the basis of an independent test set comprising the 2,500 states of the data base not used for its building, yielding an overall error rate of 4.3% (the proportion of erroneous classifications of test states of each subtree are depicted in the lower part of each node-box in Fig. 5). Out of the 1,622 fairly unstable states, only 30 are classified as stable yielding 1.85% “dangerous” errors. On the other hand, 23 marginally unstable states are classified stable, leading to small non-detection errors. There are also 52 false alarms, i.e. stable test states classified unstable by the tree.

To improve accuracy, the same data base was further exploited by building a multilayer perceptron (with a single hidden layer of 20 neurons) on the basis of the same 10,000 learning states, leading to a reduced test set error rate of 2.4%.

In terms of computational requirements we mention the following CPU times determined on a SUN Sparc10 workstation : about 1 week for the data base generation, 1 hour for the decision tree building and 1 second for testing; and 60 hours for the learning of the MLP weights and 10 seconds for testing.

Further investigations were made, concerning the data base decomposition into various topology classes for which simpler and more interpretable trees were obtained. These and other recent research results, e.g. concerning a nearest neighbor technique optimized using genetic algorithms, are described in [14].

4.2.2 Voltage security

A second rather extensive feasibility study was carried out for voltage security, on a test problem concerning the Brittany region of the EDF system. Both preventive security assessment and emergency state detection were considered [4, 15]. Figure 6 depicts in its left part the one-line diagram of the related part of the EDF system. Its subregions correspond to voltage coherent load areas, determined with respect to behavior of HV voltage magnitudes just after the loss of a generator in Plant No. 1. These regions were automatically determined in a preliminary study by non-supervised learning using a Kohonen feature map [16].

The independent variables used during the random sampling of the pre-disturbance states concerned the following : topology (single or double line or transformer outages); regional load level, unit commitment and generation dispatch; reactive support by synchronous condensers and gas turbines. To account for uncertainties the following quantities were also randomized : secondary voltage control set-points; individual HV load distribution and power factors; MV shunt compensation; voltage sensitivities of the active and reactive load powers.

A total of 13,513 random variants were drawn to yield 5000 pre-disturbance states. (The remaining 8,513 variants led to power flow computation divergence or non-convergence.) For each state about 200 attributes were computed, corresponding to key variables such as topological indicators, important EHV power flows, 400kV voltages, numbers of units in operation in power plants, total load demand, reactive shunt compensation reserves in the study region, and reactive generation reserves.

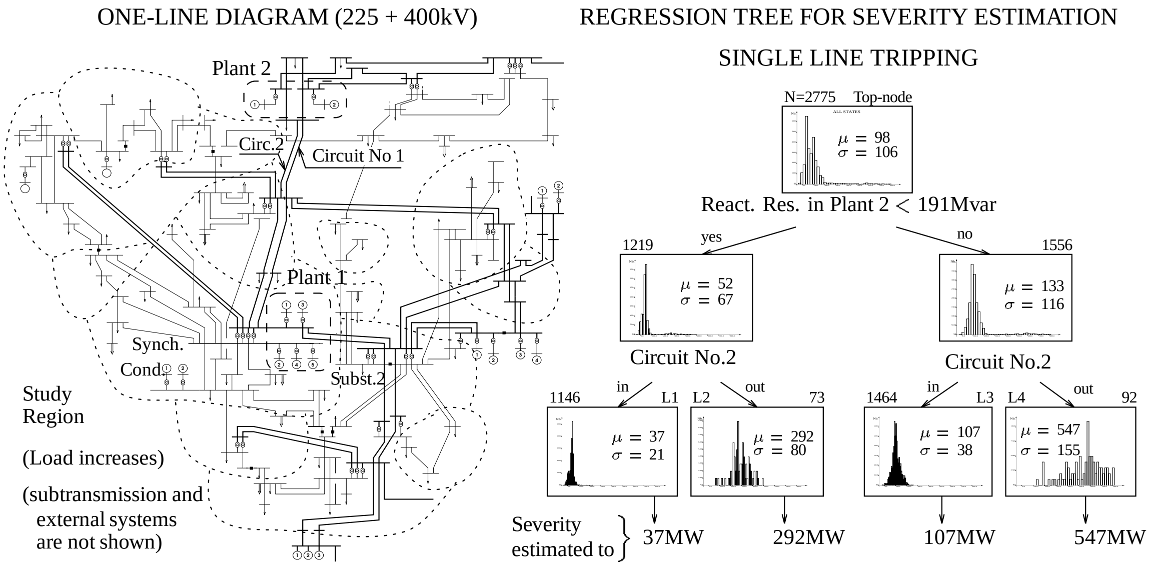


Figure 6: Voltage security assessment of the Brittany system

All in all 26 different contingencies were considered in this broad study, corresponding to synchronous condenser, generator or line tripping and busbar faults. The electrical static and dynamic models, the voltage security criterion and the load power margin computation procedure used are described in [15]. The severity of a disturbance was determined by the difference between pre- and post-disturbance load power margins in the Brittany region. Thus, in addition to the pre-disturbance margin, the corresponding 26 post-disturbance margins were computed for each operating state, yielding a total number of 135,000 load power margin computations. Overall, the data base generation required about one month of CPU time on a SUN Sparc10 workstation.

Several tens of multilayer perceptrons and even more decision and/or regression trees were built, for different disturbances and both preventive security assessment and emergency state detection. In addition, various nearest neighbor classifiers were also tried out. For illustration, we will comment briefly on the regression tree depicted in the right part of Fig. 6, built to estimate the severity of the loss of Circuit No. 1 of an important 400kV line (see the one-line diagram in Fig. 6). Each node of the tree is represented by a box containing a graphical representation of the distribution of values of the contingency severity in the learning set at this node, together with its sample mean value and standard deviation, and the number of its learning states : at the top-node, $N = 2775$ corresponds to the total number of learning states used to build the tree.

The total pre-disturbance reactive reserve available in Plant 2 is automatically selected as the best test attribute at the top-node with a threshold of 191Mvar. The learning set is split in two subsets, corresponding respectively to 1219 and 1556 states. This reduces the variance from $106^2 = 11236$ at the top-node to a mean value of $\frac{1219}{2775}67^2 + \frac{1556}{2775}116^2 = 9517$ at its successors.

Proceeding at both successors, we see that the selected test consists of checking whether Circuit 2 is in operation or not, which allows us to further reduce significantly the overall variance to a mean value of $\frac{1146}{2775}21^2 + \frac{73}{2775}80^2 + \frac{1464}{2775}38^2 + \frac{92}{2775}155^2 = 1817$. Thus, the regression tree explains $100 \times (1 - \frac{1817}{11236}) = 84\%$ of the variance of the severity.

Once the tree has been constructed, it may be used to estimate the contingency severity of an unknown state : to this end, we direct the state from the top-node to the appropriate successor according to its reactive reserve and further to a terminal node according to the status of Circuit 2. There, the severity is estimated by the mean severity of the corresponding learning states.

This very simple tree provides actually a very accurate estimate of the severity of the disturbance. Admittedly, it might be further improved by further developing some of its terminal nodes using other attributes carrying complementary information, Applying it to a representative independent test sample, the difference between this estimate and the "actual" pre-computed severity yields an overall mean error of -0.5MW and standard deviation of 43.6MW, which is indeed almost negligible if compared to the overall load level of the study region which varies between 5,000MW and 7,700MW.

5 Conclusions

In this paper we have attempted to survey state of the art and future potentials of *machine learning approaches* for *power system security assessment*.

We have described the high diversity of power system security problems so as to justify the combined use of machine learning and other statistical and neural network based automatic learning methods, in a tool box fashion. To provide insight into the possible complementary uses of these various methods, we have put the emphasis on illustrations and discussions, rather than on theoretical presentations. To render credible machine learning approaches to power system security, we have reported a small subset of results obtained with two different real-life problems.

One of the messages we would like to convey is that to make automatic learning methods really successful it is important to include the human expert in the process of deriving security information. For example, to guide the security studies it is necessary to exploit his prior expertise and then to allow him to criticize, assimilate and accept the new information. The results must therefore be provided in a form compatible with his own way of thinking. In the general class of automatic learning approaches, machine learning is presently the only one able to meet this requirement; it is therefore a key element of the tool box.

Clearly, machine learning as well as other learning methods can produce interesting security information only when they exploit representative data bases. To obtain them, the initial investment is quite important for each new security problem, but the subsequent data base generations take full advantage of the previous ones. To further enhance the approach, powerful parallel simulation environments could be developed to enable a transparent allocation of simulations on virtual machines composed of the large numbers of elementary workstations available through local or wide area networks, and not fully exploited today.

After eight years of research, we deem that automatic learning methods are indeed able to provide interesting security information for various physical problems and practical contexts. Actually, in their philosophy they are quite similar to existing practices in power system security studies, where limits are derived from simulations, though in a manual fashion. But automatic learning approaches are more systematic, easier to handle and master, in short more reliable and powerful.

These possibilities open up new perspectives to power system engineers to respond to the challenge of planning and operating future power systems with an acceptable level of security, in spite of growing complexity and level of uncertainties (e.g. due to the de-regulation of transmission systems and faster technological changes) and increasing economical and environmental pressures.

Acknowledgments

Our research was supported by several collaborations with Electricité de France and Hydro-Québec, and we are pleased to acknowledge the valuable discussions and suggestions of the engineers of the R&D division of Electricité de France and of the Operations department of Hydro-Québec.

References

[1] **DyLiacco, T. E.** *Control of Power Systems via the Multi-Level Concept.* PhD thesis, Sys. Res. Center, Case Western Reserve Univ., 1968. Rep. SRC-68-19.

[2] **Pao, Y. H., DyLiacco, T. E., and Bozma, I.** ‘Acquiring a qualitative understanding of system behavior through AI inductive inference’. In *Procs. of the IFAC Symp. on Electric Energy Systems*, pp 35–41, 1985.

[3] **Wehenkel, L. and Pavella, M.** ‘Decision tree approach to power system security assessment’. *Int. J. of Elec. Power and Energy Syst.* Vol 15 No 1, pp 13–36, 1993.

[4] **Wehenkel, L., Van Cutsem, T., Pavella, M., Jacquemart, Y., Heilbronn, B., and Pruvot, P.** ‘Machine learning, neural networks and statistical pattern recognition for voltage security : a comparative study’. *Engineering Intelligent Systems for Electrical Engineering and Communications* Vol 2 No 4, pp 233–245, Dec. 1994.

[5] **Fink, L. H. and Carlsen, K.** ‘Operating under stress and strain’. *IEEE Spectrum* Vol ??, pp 48–53, Mar. 1978.

[6] **Wehenkel, L. and Pavella, M.** ‘Advances in decision trees applied to power system security assessment’. In *Proc. of APSCOM-93, IEE Int. conf. on advances in power system Control, Operation and Management (Invited)*, pp 47–53, Dec. 1993.

[7] **Weiss, S.M. and Kulikowski, C.A.** *Computer systems that learn.* Morgan Kaufmann, USA, 1991.

[8] **Quinlan, J. R.** Learning efficient classification procedures and their application to chess endgames. In Michalski, R. S., Carbonell, J., and Mitchell, T., editors, *Machine Learning : An artificial intelligence approach.*, chapter 15, pp 463–482. Morgan Kaufman, 1983.

[9] **Boyen, X., Wehenkel, L., and Pavella, M.** ‘Fuzzy decision tree induction for power system security assessment’. *To appear in Procs. of SIPOWER’95, IFAC Symp. on Control of Power Plants and Power Systems*, Mexico, Dec 1995.

[10] **Hertz, J., Krogh, A., and Palmer, R. G.** *Introduction to the theory of neural computation.* Addison Wesley, 1991.

[11] **Duda, R. O. and Hart, P. E.** *Pattern classification and scene analysis.* John Wiley and Sons, 1973.

[12] **Kohonen, T.** ‘The self-organizing map’. *Proceedings of the IEEE* Vol 78 No 9, pp 1464–1480, Sept. 1990.

[13] **Wehenkel, L., Pavella, M., Euxibie, E., and Heilbronn, B.** ‘Decision tree based transient stability method - a case study’. *IEEE Trans. on Power Syst.* Vol PWRS-9 No 1, pp 459–469, 1994.

[14] **Wehenkel, L., Houben, I., Pavella, M., Riverin, L., and Versailles, G.** ‘Automatic learning approaches for on-line transient stability preventive control of the Hydro-Québec system’. *To appear in Procs. of SIPOWER’95, IFAC Symp. on Control of Power Plants and Power Systems*, Mexico, Dec 1995.

[15] **Wehenkel, L.** ‘Contingency severity assessment for voltage security using non-parametric regression techniques’. Paper # 95 WM 162-8 PWRS presented at the IEEE PES Winter Meeting, to appear in *IEEE Trans. on Power Syst.*, Jan. 1995.

[16] **Wehenkel, L.** ‘A statistical approach to the identification of electrical regions in power systems’. Submitted for presentation at Stockholm Power Tech, June 1995.