Proceedings of PSCC99, Volume 1, pages 324-330
The 13th Power System Computation Conference, Trondheim, Norway, June 28 - July 2nd, 1999

DATA MINING TOOLS AND APPLICATIONS IN POWER SYSTEM ENGINEERING

Cristina OLARU, Pierre GEURTS, Louis WEHENKEL
Department of Electrical Engineering - University of Liège
Sart-Tilman B 28 B-4000 Liège - BELGIUM
olaru@montefiore.ulg.ac.be

Keywords: data mining, machine learning, dynamic security assessment

Abstract: The power system field is presently facing an explosive growth of data. The data mining (DM) approach provides tools for making explicit some implicit subtle structure in data. Applying data mining to power system engineering is an iterative and interactive process, requiring an acquainted user with the application specifics. The paper describes data mining tools like statistical methods, visualization, machine learning and neural networks, exemplifying by results obtained with a DM software developed for dynamic security assessment studies. Power system engineering applications where data mining would be useful are reviewed in the second part of the paper.

1 INTRODUCTION

Like many other application areas, the power system field is presently facing an explosive growth of data. In power systems, irrespectively of the particular application, there are three main sources of data: (i) field data, collected by various devices distributed throughout the system, such as digital records; (ii) centralized data archives, such as those maintained by control center SCADA systems; (iii) data from simulations, carried out in planning or operation environments. Trends in technology (e.g. in digital protection and recording, control center architectures, simulation software and hardware ...) result in more and more data generation, and dramatic cost decrease of mass storage devices makes it possible to store them efficiently (terabytes - 10^{12} bytes - of data).

The massive raw material needs to be managed appropriately, preferably recycled. The idea is to process the information from the enormous stock of data we have, developing the possibilities for better ways to handle data and support future decision making.

Sometimes, the patterns to be searched for, and the models to be extracted from data are subtle, and require complex calculus and/or significant specific domain knowledge, or even worse, there are situations where one would like to search for patterns that humans are not well-suited to find [1], even if they are good experts in the field. In situations like power systems related problems where one is faced with high dimensional data sets that are not well understood on the whole, algorithms capable of synthesizing structures from data become a necessity.

Developing such algorithms, implementing them into a complete software environment, and applying them to real problems are the purposes of the so called data mining field. DM is a term calling for a general approach that is supported to varying degrees by a set of technologies: statistics, visualization, machine learning and neural networks. Classification, regression, clustering, summarization, dependency modeling, deviation detection, and temporal problems are expected to be solved by DM tools.

Why would data mining tools be useful in the power system field? There are a few DM related aspects: large scale character of power systems (thousands of state variables), temporal (from milliseconds to minutes, hours, weeks, years) and statistical nature of data, existence of a discrete (e.g. events such as topology changes or protection arming) and continuous (analog state variables) information mixture, necessity of communication with experts through means of visualization, on-line operation time restrictions for fast decision making, existence of uncertainty (noise, outliers, missing information).

This paper aims at reviewing power system engineering applications where data mining approaches would be very useful, starting first by describing specific DM methods and tools which have been developed in order to enable power systems data mining. For a detailed presentation of data mining tools and applications in power systems we invite the reader to have a look at [2] and the references therein.

2 DATA MINING

2.1 Data Mining process

Data mining [3, 4, 5] is the process of extracting valid, previously unknown, comprehensible, and useful information from large data bases and using it. It is an exploratory data analysis, trying to discover useful patterns in data that are not obvious to the data user. DM takes two forms [6]: *verification-driven data mining*, which extracts information in the process of validating a hypothesis postulated by a user and *discovery-driven data mining*, which automatically extract information novel for the user. Both approaches need an acquainted user with the application specifics. The quality of the mined information is a function of both the effectiveness of the data mining techniques used and the quality (often size) of the data base.

A data base is composed of a certain number of objects each one of which is described by a certain number of attributes. For a power system dynamic security assessment DM application for example, each object of the data base, called security scenario, essentially corresponds to a (simulated) dynamic trajectory of the studied system under some particular conditions [2]: the initial operating point (available equipments in operation, topologies ...), external disturbances (outages, load disturbances or faults) and dynamic modeling hypothesis (system parameters and automatic/manual actions behavior). Additionally to these attributes, each scenario is provided also with information concerning its security (discrete e.g. secure/insecure, numerical e.g. security margin), temporal behavior of some variables ....

Once the data base has been developed, the data mining process falls in five basic steps.

Data selection. The step consists of choosing the goal and the tools of the data mining process, identifying the data to be mined, then choosing appropriate input attributes and output information to represent the task.

Data transformation. Transformation operations include organizing data in desired ways, converting one type of data to another (e.g. from symbolic to numerical) defining new attributes, reducing the dimensionality of the data, removing noise, "outliers", normalizing, if appropriate, deciding strategies for handling missing data.

Data mining. The transformed data is subsequently mined, using one or more techniques to extract patterns of interest. The user can significantly aid the data mining method by correctly performing the preceding steps.

Result interpretation and validation. For understanding the meaning of the synthesized knowledge and its range of validity, the data mining application tests its robustness, using established estimation methods and unseen data from the data base.

Incorporation of the discovered knowledge. This consists of presenting the results to the decision maker who may check/resolve potential conflicts with previously believed or extracted knowledge and apply the new discovered patterns.

The whole DM process is iterative, interactive, and very much a trial and error activity. DM techniques are different one from another in terms of problem representation, parameters to optimize, accuracy, complexity, run time, transparency, interpretability. Visualization plays an important role - it may provide preliminary understanding of data, domain specific visualizations or can present the results of the mining techniques.

2.2 Statistical methods

Statistical data analyses are involved in many stages of the DM process: data cleansing (retaining outliers which may arise from human or measurement errors, estimating noise parameters), data selection (sampling facilities to reduce the scale of computation), data mining (model search), result evaluation (by descriptive statistical measures).

Statistics provide a number of data-analysis tools, like parametric and non-parametric statistical methods (supervised learning) and clustering methods (unsupervised learning) [2, 7]. Parametric methods, like discriminant analysis for classification or regression and projection pursuit regression technique, aim at obtaining decision rules or regression models. Distance based methods (like K nearest neighbors), the Kernel density estimation approach, histogram approach or "naive" Bayes approach are referred as non-parametric methods and they may be employed

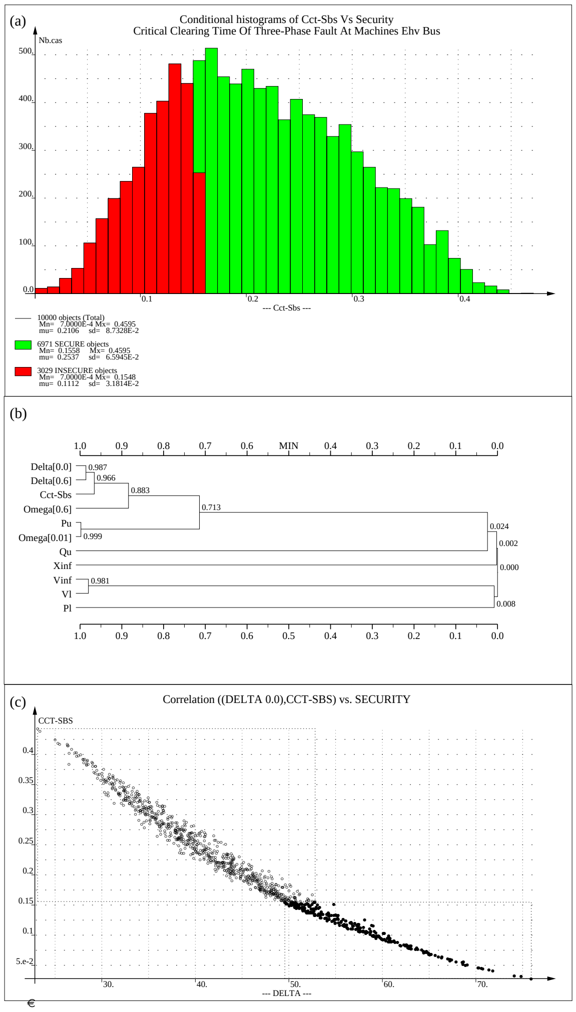

Figure 1: Example of statistical visualizations for an academic transient stability assessment problem: (a) histogram; (b) dendrogram; (c) scatter plot.

in data mining in order to find the most similar object in a data base with a given one (e.g. finding the most similar scenario in the data base with an unseen situation encountered on-line), or to estimate probability distribution for a certain numerical attribute given a set of objects. Taking a transient stability problem for example, figure 1a presents an example of one-dimensional conditional histogram built for 10.000 states of a simple One-Machine-Infinite-Bus (OMIB) system temporal data base, considering the critical clearing time (CCT) attribute space and given a three-phase short-circuit in the EHV substation normally cleared after 155ms. The system is insecure if the CCT is smaller than 155ms (dark grey on the picture) and secure otherwise.

Hierarchical attribute clustering approach (dendrogram approach), the algorithm of dynamic clusters (K-Means, ISODATA) and mixture distribution fitting algorithm (AutoClass) fall into the class of unsupervised learning statistical methods and support data mining operations like evaluating similarities among data (e.g. among security scenarios), building a set of representative prototypes, analyzing correlations between attributes. Figure 1b shows a dendrogram for the same example of transient stability, in fact a correlation analysis between some numerical attributes: CCT, active (P_u) and reactive (Q_u) power of the generator, transient machine rotor angle ($delta[t]$) and speed ($omega[t]$), voltages, reactance. Looking at such a correlation analysis one may decide which attributes to select for the next step of data mining, keeping only those which are highly correlated with the output and which, in the same time, present low mutual correlation.

Another part of DM operations involving statistics is the extraction of basic statistical summarizations:
- in numerical forms - simple descriptive statistical measures (see figure 1a) like mean (for location), standard deviation, minimum, maximum (for spread);
- in graphical forms like scatter plots or cumulative distribution graphics. Figure 1c shows a scatter plot, a visualization tool providing information on the relation between two numerical attributes (the transient machine rotor angle at time zero and the CCT) and a discrete attribute (security levels).

2.3 Visualization techniques

Because we find it difficult to emulate human intuition and decision-making on a machine, the idea is to transform the derived knowledge into a format that is easy for humans to digest, such as images or graphs. Then, we may rely on the speed and capability of the human user visual system to spot what is interesting, at every step of the data mining process: preliminary representation of data, domain specific visualization or result presentation.

Power system specific data mining tools for visualization are mainly using two-dimensional maps, colors in order to highlight groups of data which share a common property (same voltage level, same island, same status on/off ...), numerical information (voltage magnitudes, power flows, percentages ...). In this context of a large scale power system visualization [8], graphical maps should provide easy to use zooming, focusing, object and group oriented searching, navigation mechanisms between different levels of the information display, incremental editing facilities in case of configuration changes, graphical selection possibilities, simultaneous visualizations for reasons of comparison analysis, time skipping facilities, if appropriate.

All the figures provided in the paper are examples of data mining visualization tools. They are incorporated in a data mining software developed in the context of applying automatic learning techniques to power system dynamic security assessment.

2.4 Machine Learning

Machine learning (ML) [2, 4] is the center of the data mining concept, due to its capabilities to gain physical insight into the problem, and participates directly in data selection and model search steps. Assessing problems like classification (crisp and fuzzy decision trees), regression (regression tree), time-dependent prediction (temporal decision tree), ML field is basically concerned with the automatic design of if-then rules similar to those used by human experts. Decision tree induction, the most known ML framework, was found to be able to handle large scale problems due to the computational efficiency, providing interpretable results and in particular, able to identify the most representative attributes for a given task.

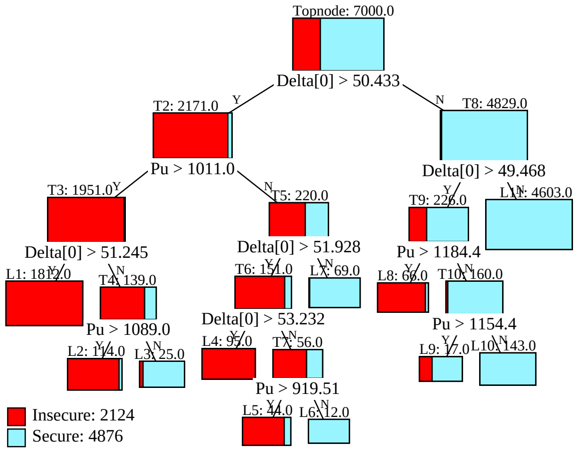

Figure 2: Decision tree for OMIB transient stability assessment.

2.4.1 Basic Decision Trees

Figure 2 shows a decision tree technique applied to the same transient stability OMIB temporal data base. The transient machine rotor angle and active power of the generator attributes where chosen to explain the output classes defined on the CCT basis: secure and insecure.

In a decision tree, a test built upon one attribute at a time is attached to each interior node, and each terminal node (leaf) is labeled with an output class. To classify a new scenario, one starts at the top node and applies sequentially the dichotomous tests encountered to select the appropriate successor. Finally, a unique terminal node is reached and the class stored there is assigned to this scenario. Note that the tree may be translated into an equivalent set of mutually exclusive rules, each one corresponding to a path from the top node to a terminal node.

For numerical attributes, starting with a given learning set, each step of the tree growing algorithm consists in selecting an attribute together with a threshold [2], in order to split the input space into *non-overlapping* subregions of objects which have the same class. Figure 2 presents every node of the resulting tree by a box area proportional to the size of the learning subset corresponding to this node (the exact size of this subset together with the name of the node are indicated above the box) and the horizontal division of each box shows the proportion of secure versus insecure states in this subset. The test used to divide a node is written under the box and each arc leading to a successor is labeled with a possible answer to this test (here, Yes or No).

To determine the generalization capabilities, the tree is tested on a test set of 3000 states different from the 7000 learning states, yielding only 7 states declared erroneously secure and 17 false alarms, i.e. an error rate of 0.8%.

"Oblique" decision trees are refined decision trees using linear combinations of attributes in its test nodes. Regression trees are extensions of the induction technique, predicting a continuous numerical output, rather than a discrete class (e.g. the CCT).

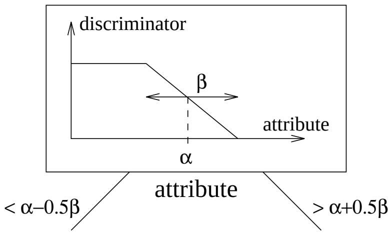

Figure 3: A test node in a fuzzy decision tree.

2.4.2 Fuzzy Decision Trees

Fuzzy decision tree technique is quite useful in the context of a data base with imprecise information from uncertain measurements or human expert linguistic knowledge translation. The tool exploits fuzzy attributes and/or fuzzy classes providing input - output mappings in the form of membership degrees. A class definition like unstable, stable, very stable in the form of a fuzzy set is more rich and valorous for an operator, than the YES/NO type of answer of the crisp classification.

Looking to a classification problem in the context of continuous numerical attributes [9], a binary fuzzy tree growing identifies the test in a node such as to define a transition region between the left and right successors. This supposes the identification of two thresholds in the domain of the chosen attribute, not only one, like for crisp decision trees. The input space is split in *overlapped* subregions of objects. It is in fact a fuzzy discretization of the attributes, based on a priori chosen shape of a "discriminator" function: gaussian, sigmoid, or piecewise linear. Figure 3 shows a test node in a fuzzy tree: α and β are the parameters to be identified, corresponding to the two attribute thresholds: $\alpha - 0.5\beta$ and $\alpha + 0.5\beta$.

Any new object is propagated through the tree in parallel, along *numerous* decision paths, and it reaches multiple terminal nodes which are then aggregated to obtain the final estimated membership degree to the target class. This causes increased computational burden, but in the same time it offers continuity of the output with respect to inputs, fact which permits a global parameter optimization process for the final tree, thus, a better tree accuracy.

2.4.3 Temporal Decision Trees

Temporal decision trees [10] are a first and recent attempt to adapt the decision tree induction method to the treatment of temporal data. Previously, the only way to incorporate temporal data into a basic decision tree was to sample the input signals at one or several particular moments and to use these new informations as inputs to the tree building process (see fig. 2). Temporal trees try to get rid of this temporal attributes sampling, by taking into account explicitly the role of time during the induction process.

Figure 4 shows a simple temporal tree built from a data base generated by numerical simulation for a large scale power system blackout study. This tree aims at detecting voltage collapses in a particular region of the system by means of generator excitation currents ("IF-gen$i$" denotes excitation current of the generator i) and voltage magnitudes at different buses ("U-bus$j$" denotes voltage magnitude at bus j).

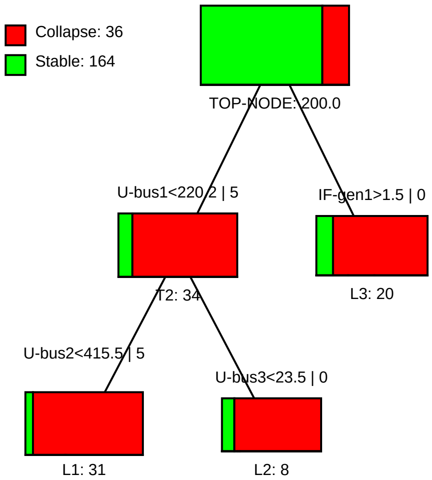

Figure 4: Temporal decision tree for voltage collapses detection.

In a temporal tree, a test (which is now a time function) is attached to each arc and any new scenario is propagated through the tree in parallel, along *all* the paths, starting at the root, until a condition along the path is false, or a terminal node is reached. If the scenario reaches at least one terminal node, it is classified insecure, otherwise it remains classified as a secure scenario. Unlike in the case of decision trees, here a scenario may go deeper in the tree only when a test is true, and it may reach multiple terminal nodes at the same time or may never reach any of them. Particularly, a test in fig. 4 has the form "attribute < threshold (resp. >) | temporization". A test like this is true at one specific moment if attribute values were below (resp. above) the given threshold during at least the temporization.

When building a temporal tree, the method takes into account two contradicting requirements (by means of a user parameter): on one hand, the discrimination capability of the tree (in terms of false alarms and non-detections) and on the other hand, the degree of anticipation (the time needed to detect). This allows the user to build different detection rules for different purposes (very conservative versus very anticipative rules).

Although this method has of course not yet reached the maturity of decision trees, the first results look very promising and prove the usefulness of such a method in the context of power systems engineering.

2.5 Artificial Neural Networks

Artificial neural networks (ANN) are recognized in the automatic learning framework as "universal approximators", with massively parallel computing character and good generalization capabilities, but also as black-boxes due to the difficulty to obtain insight into the relationship learned [11, 12].

They are used within the data mining step:

- to generate a regression model that can predict future behavior, on the basis of a data base with input-output pairs of continuous numerical historical information (supervised learning). The network acts like a function associating numerical outputs to any new object of known attribute values.

• to automatically represent a data set by a small number of representative prototypes, preserving the topological properties of the original attribute space (unsupervised learning). The output corresponding to a new object is defined as the nearest prototype in the attribute space [13, 14, 15].

The first approach is used in power systems for modeling practices (e.g. load forecast, state estimation, severity estimation), the second one for graphical data inspection, visualization, correlation analysis (e.g. searching similarities between operating states or between physical parameters).

3 DATA MINING APPLICATIONS

Engineers trying to solve power system related problems should look at the whole tool-box of data mining methods, and not hesitate to combine different techniques to yield a full, practical solution. Data selection step may be performed with a decision tree, a clustering approach, or a correlation analysis and later on, the result may serve as input for other supervised techniques, possibly with the problem decomposed into simpler subproblems.

3.1 Dynamic Security Assessment

Power system security assessment has the special feature that the data bases are generated by computer based simulation, and this because the actual experience of instabilities is fortunately very small for most power systems. This is a laborious task but in the same time an advantage, being possible to generate as many scenarios as are deemed necessary to solve a particular task.

Possible uses of data mining to security assessment would be:

- providing continuous security margins for a given scenario: CCTs, load power margins, severities, global stability limits;
- identifying the relevant attributes for the security class prediction of a given power system;
- determining if an operating state is stable or not with respect to a given fault, or to a set of contingencies;
- determining which operating points lead to a voltage collapse for a given contingency;
- determining for a given operating state the faults likely to create instability;
- contingencies ranking in terms of their severity for a range of operating conditions;
- identifying conditions characterizing pre-fault attributes of stable operating states for a given set of possible faults;
- partitioning all the possible scenarios into families of dynamic behaviors: stable and failure modes; determining criteria to identify failure modes in real-time;
- partitioning the power system map into voltage coherent behavior regions;
- searching for a specific kind of scenario;
- early detection of voltage risks for a given scenario;
- inferring means for control if an incipient instability is detected early enough;
- extracting (fuzzy) if-then rules from all the problems solved by means of machine learning.

3.2 Controller and Protection Design

In power systems, design problems mainly concern the identification of appropriate locations for some devices, their dimensioning and the tuning of some of their parameters. Typical applications are: study concerning a new static VAr compensator (SVC) (the best geographical location, type and capacity, possibly parameters tuning), re-tuning the parameters of power system stabilizers equipping some plants, selecting the number and location for new measurement devices, defining pilot nodes for secondary voltage control systems, tuning a transformer tap-changer blocking device.

There is a variety of control problems, concerning centralized/local controllers operating at different time-frames and acting on different physical variables. These controllers need to be robust with respect to changing characteristics of a power system, due to variation of load level and topology. Since the latter are generally non-linear it is necessary to use simulation tools in order to validate controller robustness and sometimes tune them. Data mining can be used in order to make such studies more systematically, taking better advantage of simulations.

A DM process applied to the above SVC design problem should start with the identification of the electrically "coherent" regions in the system. A small number of voltage coherent regions may be found out by means of unsupervised learning, analyzing correlations among voltages at different nodes of the system. Then, a more refined analysis of the individual voltages in each region would allow to identify "central" buses where the overall behavior of the zone is best observed. To study the best substation to place the SVC and to find out which parameters of the operating point would influence most strongly the behavior of the device, a new data base is generated by random sampling different possible locations, operating conditions, assumptions and modelings (steady state, small signal, large disturbances), and various data mining tools may be implied. In order to finalize the design, a third more refined study could be carried out for fixing SVC parameters: several possible sizes and parameters could be screened out.

Adaptive (control and protection) systems are an alternative to robust designs, in order to cope with the variety of operating conditions under which such systems must operate. Data mining would allow one to determine the parameters to which the optimal control law of a plant is most sensitive and then learn the scheduling rules to select the appropriate control law in real-time.

3.3 Modeling

In order to carry out simulations, be it for security assessment or design, it is of course necessary to build up models; while many components in the power system may be modeled rather accurately, by choosing model structure from first principles and using classical identification techniques to determine parameters, some other parts are intrinsically difficult to model, either because they need to be represented in a reduced way (this is typically the case of load models and external system models), or because they change with time (due to aging and/or modifications carried out during maintenance operations) or they vary significantly from one geographical region to another (due to weather conditions). We believe that data mining could be a very useful tool in order to improve the latter type of models, and also to monitor the former ones in a systematic fashion in order to develop better models.

Data to be mined can be available in this case by recording information from real-time measurement systems and disturbance records. By data mining tools, it would be possible to estimate how accurately their derived model approximates real-life, i.e. how much uncertainty remains which is not explained by the model. And this information may be used in the context of problems using data base generated by means of these models. For example, a certain standard deviation of the security margin in a generated data base, due to uncertainty in the modeled load distribution, translates into inevitable classification errors when trying to assess the security class of a scenario. Then we know that, for minimizing these errors we should try to model more accurately the load, rather than to further search for improving data mining methods.

3.4 Estimation, Identification, Forecasting

In order to build up and maintain statistical models about load patterns, measurement and signaling errors, data gathered in control centers archives may be exploited by data mining tools.

Short-term load forecasting is one of the areas very investigated from the point of view of neural networks application (supervised and unsupervised). In this context, a data base may be easily updated to contain past load demands and past data about some typical factors: local/global weather conditions (temperatures, winds, humidity), type of the hour, type of the day, tariff strategies, social events like national days off, strikes. Mining this type of data gives a better understanding about what parameters are more important for the load evolution, how different geographical regions act from the point of view of weather conditions and how this influences the load, and accurately predict the load demand for a given new situation [16].

An interesting area for research concerns very short term *nodal* load forecasting, to be used in enhancing real-time system monitoring and security assessment. Such predicted values would enrich the information exploited by state estimators, providing the required redundancy to detect topological errors. Further, obtaining a 30-minute prediction of the individual load demands would permit security assessment on predicted state rather than on past information.

3.5 Monte-Carlo based Planning

In the off-line planning environment, engineers use Monte-Carlo simulations in order to estimate expected values of quantities such as operating costs and load-curtailment for future operating conditions. It consists of using random sampling to screen possible scenarios according to their probability, together with numerical methods to compute for each scenario the operating cost and amount of load curtailed, and averaging the values for the simulated scenarios. Presently, most tools used within this context compute only global sample statistics (averages, standard deviations, quantiles) of important quantities chosen by the engineer, but throw away the details concerning each individual sample.

Preliminary investigations show that much richer information can be extracted by storing the results from individual samples (simulated operating scenarios) in a data base, and further analyzing the latter by data mining. For example, this approach allows one to detect outliers (which are often responsible of high variance and bias in the estimates) and lets the engineer judge whether or not they must be taken into account. Also, by carrying out systematic studies of relationships (correlations) between different output parameters, better insight into the physical structure of costs and other decision variables can be gained, which eventually will lead to better strategic decisions.

Data mining can also be used in order to compare the distributions of simulated scenarios and those encountered in real life, which in turn leads to improvements in the probabilistic models used in the Monte-Carlo simulations. Furthermore, by carrying out the analysis of variance of output parameters, it is possible to find out in which region of the sampling space variance is higher, and thereby to define more efficient sampling schemes. In the same way, regression techniques may be used in order to build automatically approximate input/output models which can then be used as control variates in order to further reduce variance [17, 18].

3.6 Monitoring

Monitoring is a very broad application field, ranging from the monitoring of the individual devices and plants, to sophisticated system monitoring. The purpose is to detect as early as possible when a system departs from its normal conditions. It includes many of the activities of a control-room, in particular security assessment. Power plant monitoring (burner, reactor, dam), transformer monitoring and system oscillations monitoring are applications which can make use of data mining.

In particular, temporal decision trees are an appropriate tool for monitoring, given the dynamic time-varying nature of attributes and the ability of temporal trees to early predict modes of bad behavior (e.g. fast electromechanical transients, mid-term voltage collapses, cascades of overloaded line tripping...).

Some of these applications have already received attention from researchers, but many possibilities still remain to be explored. In particular, so as to respond to present unbundling trends, future system monitoring will mostly rely on hierarchical approaches. In this hierarchy, the lower levels will correspond to local monitoring of plants or sets of plants of individual generating companies or load aggregators, and the higher levels will correspond to security and transaction monitoring at the system level. At the lower levels, the problem will be to monitor system changes which may impact the functioning of local plants, from local information and possibly global status information obtained from the upper level monitoring system. At the upper levels, the problem will be to use global information together with outputs of local monitoring systems in order to provide an overall picture of the system, and identify those abnormal states which cannot be identified from local information only. Which information to share between lower and upper levels, and how to make the best use of it will be the main questions to solve. We believe that probabilistic simulations, together with data mining and other intelligent information processing approaches may provide valuable tools in this context.

4 CONCLUSION

This paper presented power system areas where data mining may be systematically used for extracting information, exploiting data already available from field measurements, computations and centralized archives, building up or improving models, early predicting abnormal scenarios.

Of course, it is useless to apply data mining to an invalid data base with high measurement or estimation data errors, or to try to precisely estimate numerical outputs which present a priori high variance. A data mining tool explains as much information as it is stored in the data which is mined (a derived model is strongly dependent on the learning set used).

An important part of data mining result errors are due to uncertainties in modeling and generation of scenarios in a data base in discordance with the real probabilities of system elements outages. That is why the errors do not have a meaning by themselves, they just provide a practical means to compare efficiencies of different criteria applied to the same data base.

There are three dimensions along which data mining may complement classical system theory oriented methods for power systems:

- **Computational efficiency.** By using synthetic information extracted by DM, instead of numerical methods, much higher speed may be reached for real-time decision making. Further, in terms of data requirements, DM may require only significant and/or available input parameters data base, instead of a "full" description of the system model.

- **Anticipative physical insight.** The present day practice generally handles new problems *after* that some undesirable consequences have already been observed on the system. Carrying out DM studies will allow the engineers to have a more anticipative view on potential problems.

- **Management of uncertainties.** The behavior of a power system will always have some unexpected experiences (e.g. a relay which mis-operated, an operator which did not behave as expected, a set-point which was different from prescriptions, a load which were modeled inappropriately ...). DM copes with this problem by making use of more simulations carried out by relaxing assumptions on the dynamic models used.

The coupling of two factors - the immediate needs and the difficulty of the problems to be solved - promises to keep the data mining field application in power systems on a healthy development path.

References

[1] U. M. Fayyad, "Data Mining and Knowledge Discovery: Making Sense Out of Data", IEEE Expert, Intelligent Systems and their Applications, pp 20-25, October 1996.

[2] L. Wehenkel, "Automatic learning techniques in power systems", Boston, Kluwer Academic, 1998.

[3] A. A. Freitas and S. H. Lavington, "Mining very large databases with parallel processing", Boston, Kluwer Academic, 1998.

[4] R. S. Michalski, I Bratko and M. Kubat, "Machine Learning and Data Mining. Methods and Applications", Chichester, Willey, 1998.

[5] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy (editors), "Advances in knowledge discovery and data mining", Menlo Park / Cambridge, AAAI Press / MIT Press, 1996.

[6] E. Simoudis, "Reality check for data mining", IEEE Expert, Intelligent Systems and their Applications, pp 26-33, October 1996.

[7] V. Cherkassky and F. Mulier, "Learning from data. Concepts, Theory, and Methods", New York, Willey, 1998.

[8] P. Geurts and L. Wehenkel, "Visualizing Dynamic Power System Scenarios for Data Mining", Proc. of LESCOPE'98, Large Engineering Syst. Conf. on Power Engineering, pp 217-224, Halifax, July 1998.

[9] X. Boyen and L. Wehenkel, "Automatic induction of fuzzy decision trees and its applications to power system security assessment", Fuzzy Sets and Systems, Vol. 102/1, pp 3-19, February 1999.

[10] P. Geurts and L. Wehenkel, "Early prediction of electric power system blackouts by temporal machine learning", Proc. of ICML-AAAI'98 Workshop on AI Approaches to Time-series Analysis, Madison, pp 21-28, July 1998.

[11] S. Haykin, "Neural Networks. A comprehensive foundation", New York, IEEE Press, 1994.

[12] C. M. Bishop, "Neural Networks for Pattern Recognition", Oxford, Clarendon Press, 1995.

[13] H. Mori and Y. Tamura, "An artificial neural-net based approach to power system voltage stability", Proc. of the 2nd Int. Workshop on Bulk Power System Voltage Phenomena - Voltage Stability and Security, Deep Creek Lake, pp 347-358, August 1991.

[14] D. Niebur and A. Germond, "Power system static security assessment using the Kohonen neural network classifier", Proc. of the IEEE Power Industry Computer Application Conference, pp 270-277, May 1991.

[15] L. Wehenkel and Y. Jacquemart, "Use of Kohonen feature maps for the analysis of voltage security related electrical distances", Proc. of ICANN'95, Int. Conf. on Artificial Neural Networks, NEURONÎMES'95 (Industrial conf.), Paris, pp 8.3.1-8.3.7, October 1995.

[16] T. S. Dillon and S. Sestito, "Short Term Load Forecasting using Neural Networks", in Neural Networks Application in Power Systems edited by T. S. Dillon and D. Niebur, pp 387-404, London, CRL Publishing, 1996.

[17] A. Breipohl, F. N. Lee, J. Huang, and Q. Feng, "Sample size reduction in stochastic production simulation", IEEE Trans. on Power Systems, 5(3), pp 984–992, August 1990.

[18] M. V. F. Pereira, M. E. P. Maceira, G. C. Oliveira, and L. M. V. G. Pinto, "Combining analytical models and Monte-Carlo techniques in probabilistic power system analysis", IEEE Trans. on Power Systems, Vol. 7/1, pp 265–272, February 1992.