International Journal of Smart Grid and Clean Energy

An overview of decision tree applied to power systems

Chengxi Liu[a*], Zakir Hussain Rather[a,b], Zhe Chen[a], Claus Leth Bak[a]

[a] *Department of Energy Technology, Aalborg University, 9220 Aalborg Denmark*
[b] *Innovation Centre, kk-electronic, a/s, 9220 Aalborg Denmark*

Abstract

The corrosive volume of available data in electric power systems motivates the adoption of data mining techniques in the emerging field of power system data analytics. The mainstream of data mining algorithm applied to power system, decision tree (DT), also named as classification and regression tree (CART), has gained increasing interests because of its high performance in terms of computational efficiency, uncertainty manageability, and interpretability. The fundamental knowledge of CART algorithm is introduced in this paper, followed by an overview of a variety of DT applications to power systems for better interfacing power systems with data analytics.

Keywords: Classification and regression tree, data mining, decision tree, power system data analytics

1. Introduction

With the development of information and communication techniques, conventional power grids are being modernized, with improved efficiency, reliability, economics and sustainability of generation and distribution of electricity. Currently, power system data are not only being gathered by centralized monitoring systems, e.g. WAMS, SCADA/EMS systems for online operation, control and protection, but also by a large number of distributed devices e.g. smart-meters, fault recorders etc., for further management and analysis. In addition, data from simulations, carried out by power system analysts are also playing an important role in the environment of power system planning and operating. Therefore, modern power systems are enriched with huge quantity of online and historical data, ready for extraction, processing, analysis and exploration.

Data mining techniques, involving methods of artificial intelligence, machine learning and statistics, are computational processes of discovering the useful information of data patterns from large data sets. Among many data mining techniques, especially those with "white box" nature, such as artificial neural network, support vector machine etc., DT algorithm has gained increasing interests because it not only provides the insight information of data sets with low computational burden, but also reveals the principles learnt by DTs for further interpretation. Besides, under the background that there are more and more uncertain factors in power systems, DT algorithm is capable of managing these uncertainties by statistical methods to increase the reliability.

This paper introduces the fundamental knowledge of DT and its extended algorithms, presents state-of-the-art of DT algorithms applied to power systems, and then gives many options of software to implement DT based algorithms.

2. Principles of Decision Trees

Decision tree, also named as classification and regression tree (CART), is a decision support tool which uses a binary tree-like graph or model to reveal the hidden relationship between inputs and outputs. It was first developed by Breiman *et al.* in the 1980s [1] and was firstly introduced into the field of power systems by Wehenkel *et al.* in 1989 [2]-[5].

---

* Manuscript received July 11, 2013; revised September 11, 2013.
Corresponding author Tel.: +45 2252 8660; *E-mail address*: cli@et.aau.dk.

Fig. 1 shows the architecture of data mining applications to power system. Sufficient amount of data are collected from power system which is used as the learning database. Data mining engine plays a significant role to find out important patterns that are hidden in the learning database followed by the evaluation and interpretation of the created model. Thereafter, online data is obtained from the system and used to predict the classification or the target value of the output. Finally, actions of control or protection are suggested or automatically taken if necessary.

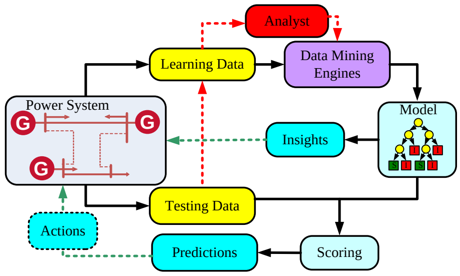

Fig. 1. Architecture of data mining in power systems.

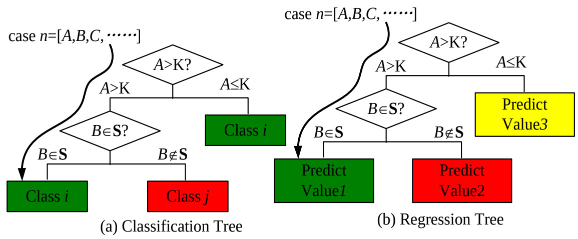

Fig. 2. A simple (a) classification tree and (b) regression tree.

As shown in Fig. 2, given a set of attributes (i.e. *A*, *B*, *C*,…) as input predictors of case *n*, the target value (i.e. Discrete or Continuous) of the case can be predicted by dropping the case downward along a path of "if-then" questions from the root node to a terminal node of a DT. The DT is classification tree (CT) if the target is a discrete class, while it is regression tree (RT) if the target is a continuous value. The vector of predictors can be composed of both numerical variables (e.g. *A*) and categorical variables (e.g. *B*). Variables are called numerical variables if their measurements are real numbers, while called categorical variable if it takes values from a finite set which may not have any natural ordering.

2.1. Accuracy estimation of decision tree

A database composed of a number of cases is necessary for training the DTs. Each case in the databases consists of a vector of predictors and a target, the predictors usually serve as the input attributes, and the target is the output. The training process of the DT is to iteratively split the dataset into 2 subsets. In CT, and the fundamental idea to select each splitting rules is such that to make the cases in each of the divided subset are as pure as possible. While in RT, the splitting rules are selected to minimize the overall absolute deviation or square error. Commonly used indices are used to estimate the cost of misclassification of a CT and the cost of error of a RT respectively, as defined in (1) and (2).

$$R_{CT}^{TS}(d)=\frac{1}{N^{TS}}\sum_{i,j}\left[c(i\mid j)\cdot X\left(d_j(x_n)\neq j_n\right)\right] \tag{1}$$

$$R_{RT}^{TS}(d)=\frac{1}{N^{TS}}\sum_{k}\left[c(k)\cdot\left|d(x_n)-y_n\right|\right] \tag{2}$$

where N^{TS} is the number of test cases, $c(i|j)$ is the cost of misclassifying a class j case as a class i case, $c(k)$ is the cost of prediction error in node k of the RT.

Usually, there are 3 methods of internal test to estimate the accuracy of a DT. The first, overoptimistic, is the resubstitution estimates, which is computed using the same data used to train the DT. The second is the test sample estimates. The cases in the database are randomly divided into a learning set (LS) and test set (TS). The LS is used to train the DT, while the TS is used to evaluate the accuracy of the created DT. The third, cross-validation, is parsimonious with data and preferred for databases with small size. The cases in the database L are randomly divided to V subsets with equal size (i.e. L_1, L_2, …, L_V), For every iteration i, L-L_i is the LS, and L_i is TS. The accuracy is the average estimate of all V iterations.

2.2. Splitting rules of decision tree

Impurity functions are adopted to find optimal splits of a CT, such as GINI index in (3), in which $p(j|t)$ is the prior probability of class j case given that the case is in node t. If split δ of a node t is predicted to send a proportion p_L of the data cases to left node t_L and proportion p_R to right node t_R, the decrease of impurity is defined by $\Delta I(\delta, t)$, as given in (4). The optimal selection of splitting rules for a node can be calculated by repeated attempts to maximize the decrease of impurity $\Delta I(\delta, t)$ so as to minimize the impurity of the whole tree, as defined in (5), in which T_t are terminal nodes of the CT. Similarly, the splitting rules of a RT are based on iterative attempts to maximize the decrease of error of RT, so (3)~(5) for CT have their counterparts for RT, as defined in (6)~(8).

$$I_{GINI}(t) = \sum_{i \neq j} p(i \mid t) p(j \mid t) = 1 - \sum_{j} p^2(j \mid t) \tag{3}$$

$$\max.\Delta I(\delta, t) = \max.I(t) - p_L(t) I(t_L) - p_R(t) I(t_R) \tag{4}$$

$$I(T) = \sum_{t \in T_t} I(t) = \sum_{t \in T_t} i(t) p(t) \tag{5}$$

$$E(t) = \frac{1}{N_t} \sum_{x_n \in t} \left(y_n - \overline{y(t)} \right)^2 \tag{6}$$

$$\max.\Delta E(\delta, t) = \max.E(t) - p_L(t) E(t_L) - p_R(t) E(t_R) \tag{7}$$

$$E(T) = \sum_{t \in T_t} E(t) = \sum_{t \in T_t} s^2(t) p(t) \tag{8}$$

2.3. Pruning of decision tree to the right size

The optimal DT is selected by the criterion of accuracy while dropping the cases in TS to the DTs. Too large a tree will have a higher misclassification rate or error than the right sized one due to the over-fitting problem, so the pruning process is needed to delete some branches or some nodes of an over-fitted tree to the right size one. The misclassification cost defined in (1) and the error cost in (2) are used to evaluate the accuracy of CT and RT respectively, so the optimal DT is the subtree with minimum misclassification rate or error rate selected from the all pruned sub-trees, as defined in (9) for CT and (10) for RT.

$$R(T_{CT}^{opt.}) = \min_{k} R_{CT}^{TS}(T_k) \tag{9}$$

$$R(T_{RT}^{opt.}) = \min_{k} R_{RT}^{TS}(T_k) \tag{10}$$

3. Applications of Decision Tree in Power Systems

3.1. Security assessment

Among the many other applications of DT in power systems, security assessment is the most versatile [2]. With the help of DT, the information regarding scenarios of power system (e.g. operating conditions, contingencies, topologies, voltage and power etc.) can be expressed as a model to assess and predict the security states of scenarios from a vector of input attributes.

DT is first introduced to assess the transient stability of a simple electric power system [3]. Later, transient stability assessment is carried out for a real power system, i.e. EDF System in France, by enhancing the reliability of DTs for seen and unseen cases [4]. In [5], new methods are proposed to reduce the probability of missed alarms by adjusting the prior probabilities and misclassification costs of unstable cases, and also to reduce the probability of false alarms by incremental tree development scheme. To evaluate the overall transient stability assessment for a large-scale power system, multi-contingency tree is first presented in a more compact and easier fashion compared with single-contingency trees.

In [6], DT is first used to assess the security of a median island power system (Lemnos Island system in Greece) with high penetration of renewable energy, taking the wind speed and solar radiation as candidate attributes. In [7], a DT based method for the optimal dispatch of primary reserve services considering security margins as constraints is tested on Crete power system in Greece.

DT is first adopted for online dynamic security assessment (DSA) for a large scale power system (Entergy system in USA) in a practical approach [8]. This scheme builds and periodically updates DTs offline to decide critical attributes as security indicators. Then DTs provide online DSA and operation guidelines based on real-time measurements from WAMS. The security of current scenario is voted by a whole path of DTs instead of just by terminal nodes to improve the robustness of DTs. In [9], DT was used especially for voltage security assessment for AEP system in USA based on the information of the past and forecast 24-h ahead operating conditions. Hourly updated DTs by including newly predicted system conditions, and multiple optimal DTs, corrective DTs are proposed for robustness improvement. In [10], a real-time security assessment tool was proposed to evaluate four post-contingency security issues of SRP in USA, i.e. voltage magnitude, thermal limit, voltage security and transient stability issues.

CT can also be combined with other algorithms for power system security assessment, e.g. fuzzy logic [11], mathematical morphology [12]. In [11], a fuzzy rule-based DT is proposed for rapid stability assessment. The voltage and rotor angle curves are proceed in both time and frequency domain to extract the area-wise and system-wise features for DT training. Then DTs can accurately predict the security of Hydro-Quebec system for decision making as early as 1sec or 2sec after fault clearing.

3.2. Preventive and corrective control

Continuously growing demand for electricity has forced modern power systems to operate closer to secure operating limits. Also, the increasing penetration of large scale renewable energy may impact power grids by bringing more uncertainties to grid operations. DT is a suitable tool to provide emergency control of power system in terms of its high efficiency, interpretability and uncertainty manageability.

In [13], DT is used to determine the reserves of Crete Island system in Greece. Then online preventive control is suggested with optimization on the cost of load shedding and the cost of spinning reserves. In [14], the application of DT for online steady state security assessment of Hellenic system in Greece is presented. The DTs extract the relevant information regarding line loading and voltage level and express them in terms of controllable values, i.e. generator bus voltage, active power generation and transformer taps. Preventive and corrective control strategies are suggested with respect to these controllable values.

In [15], integral square generator angle (ISGA) index is proposed to judge the severity of contingencies for transient stability assessment. DT is trained by a large number of time-domain simulations of contingencies to give guidelines of ISGA, so as to trigger one-shot event-based corrective control to stabilize the transient events at any time intervals of 1sec after the disturbance. In [16], one-shot response-based control is developed in response to the loss of synchronism detection.

In [17] and [18], DTs are used for preventive control (i.e. generation rescheduling) and for corrective control (i.e. load shedding) by determining the dynamic security regions. Then the preventive control and corrective control are optimized in terms of the fuel cost and the amount of load to be shed, respectively. In [19], a new methodology that using DTs in tandem is proposed for Danish power system with large scale of renewables and distributed generations, one DT about measurable variables is employed for DSA to identify potential security issues, and the other DT about controllable variables provides online decision support on preventive control strategies against those issues. In addition, DT can also be used for controlled islanding [20], [21] and load shedding [22] and so on.

3.3. Protection

Some early explorations use the DTs to trigger the protection in real time, using a short window of post-fault curves (rotor angle, speed and acceleration [23] and *R-Rdot* relay [24], [25] in progress to train a DT model, which can predict whether system insecurity/instability or even system collapse is going to occur before it really happens. Some uncertain factors in the network composed of fault duration, fault location, system operating conditions and network topology are considered to test the robustness of the created DTs, but the reliability still can not meet the standard of industrial application. In [26], the sensitivity of prediction accuracy is demonstrated for online transient stability assessment. It is found that the prediction accuracy is increasing when the power system is getting more close to the system collapse.

In [27], a methodology is proposed to reduce the likelihood of hidden failures and potential cascading failures by adjusting the security/dependability of adaptive protections. Aided with WAMS, the scheme identifies prevailing system conditions. When the system is in “safe” state, a bias toward “dependability” is desired; when the system is in “stressed” state, the protection settings is altered in favour of “security”.

3.4. Forecasting, estimation, and identification

CT and RT usually combine with unsupervised learning techniques to forecast the load consumption, to estimate the variables, and to identify the parameters in power systems.

In [28], Figueiredo *et al.* deals with the load forecast issue of EDP (Portuguese Distribution Company). Firstly, unsupervised learning (clustering) is used to obtain partitions of historical data into a set of consumer classes, then supervised learning (DT) is adopted to describe each class by rule-based classifications and create a DT model to assign consumers to the existing classes. The objective is to find the relevant knowledge about how and when consumers use electricity.

In [29], DT is used to estimate the line flows and bus voltage following an outage event in an efficient manner. The approach has been successfully tested by Taiwan system in China. In [30], a methodology is proposed for southern Spanish generation company to estimate the daily load patterns and their associated probability of non-connected unit. DT is used to recognize the load pattern so as to approximately predict when its generating units are connected to alleviate the network constraints.

RT can also combine with other algorithms. In [31], a new method based on a hybrid technology of optimal regression tree and an artificial neural network is proposed to discover the rules in short-term load forecasting. Regression tree contributes to the clustering of input data, while ANN is used to predict the 1-step-ahead load. The conventional CART regards the midpoint of two learning data as a split value. Therefore, CART misclassifies the data close to the splitting values. Therefore, Mori *et al.* combines fuzzy logic with regression tree to predict the short-term load, in which two fuzzy membership functions were assigned to express the left and right nodes so as to increase the accuracy of prediction [32].

3.5. Fault diagnosis

DT is one of the strong tools in power system fault diagnosis because of its ability to process and analyze the time-series signal. In [33], DT is used in fault diagnosis of power distribution lines, both discrete attributes (operating states) and continuous attributes (voltage and current) are included in the predictors in order to deal with the variance of fault resistance and the noise of measurement. In [34], a decision tree-based method is proposed to detect High Impedance Fault in distribution feeders using attributes of RMS and harmonics of current and voltage so as to reduce sampling rate of protection relays.

4. Extended Algorithms and Software

DT is based on a database, which is definitely significant for the quality and practicality of the DT model. There is a tradeoff between the information content in the database and the computational cost of creating a large database, so many new methods are proposed to maximize database information content and minimize computation burden. In [35], Lartin super-cube sampling is first used for efficiently identifying the security boundary region. In [36], bisection methods together with importance sampling are used to efficiently prepare the hourly database for online DSA.

As mentioned before, the algorithm of CART has advantages in terms of computational efficiency, uncertainty manageability, interpretability, as well as the ability to handle missing values [37], but it also has weakness, for instance, piece-wise sharp decision boundaries, affected by data interaction, missing the effect of weak attributes. Therefore, some extended algorithms, such as Fuzzy DT [38], ANN DT [39], [40] Random Forest [41-44], Tree Net [45] etc. are invented for strengthen the original CART algorithm.

There are two classes of software that provides DT based algorithms. One is commercialized software, e.g. SPM [46], SSPS [47], Elefant [48], Lush [49]. The other is open-source software e.g. RapidMiner [50], Shogun [51], Spider [52], Plearn [53], PyML [54], Weka [55] and MATLAB/Statistic.

5. Conclusion

The paper highlights the importance of power systems data analytics and their potential applications. Firstly, it introduces the fundamental of decision tree algorithm, followed by the latest state-of-the-art of decision tree algorithms applied to power systems. Moreover, advantages and weakness of decision tree algorithm are discussed. Finally, many choices of software to implement DT algorithms are suggested.

Acknowledgements

The authors would like to acknowledge the Danish Strategic Research Centre for providing the financial support for the project "Development of a Secure, Economic and Environmentally-friendly Modern Power Systems" (DSF 09-067255).

References

[1] Breiman L, Friedman J, Olshen RA, Stone CJ. *Classification and Regression Trees.* Belmont: Wadsworth, 1984.

[2] Wehenkel L. Machine learning approaches to power-system security assessment. *IEEE Intelligent Systems and Their Applications*, 1997, 12(5):60-72.

[3] Wehenkel L, Van Cutsem Th, Pavella M. An artificial intelligence framework for on-line transient stability assessment of electric power systems. *IEEE Trans. Power Syst.*, 1989, 4(2):789-800.

[4] Wehenkel L, Pavella M. Decision trees and transient stability of electric power stability of electric power systems. *Automatica*, 1991, 27(1):115-134.

[5] Wehenkel L, Pavella M, Euxibie E, Heilbronn B. Decision tree based transient stability method a case study. *IEEE Trans. Power Syst*., 1994, 9(1):459-469.

[6] Hatziargyriou N, Papathanassiou S, Papadopoulos M. Decision trees for fast security assessment of autonomous power systems with a large penetration from renewables. *IEEE Trans. Energy Convers.*, 1995, 10(2):315-325.

[7] Papadogiannis KA, Hatziargyriou ND. Optimal allocation of primary reserve services in energy market. *IEEE Trans. Power Syst.*., 2004, 19(1):52-659.

[8] Sun K, Likhate S, Vittal V, *et al.* An online dynamic security assessment scheme using phasor measurements and decision trees. *IEEE Trans. Power Syst.*., 2007, 22(4):1935-1943.

[9] Diao R, Sun K, Vittal V, *et al*. Decision tree-based online voltage security assessment using PMU measurements. *IEEE Trans. Power Syst.*., 2009, 24(2):832-839.

[10] Diao R, Vittal V, Logic N. Design of a real-teime security assessment tool for situational awareness enhancement in modern power systems. *IEEE Trans. Power Syst.*., 2010, 25(2):957-965.

[11] Kamwa I, Samantaray SR., Joos G. Development of rule-based classifiers for rapid stability assessment of wide-area post disturbance records. *IEEE Trans. Power Syst.*., 2009, 24(1):258-270.

[12] Khoshkhoo H, Shahrtash SM. On-line dynamic voltage instability prdiction based on decision tree supported by a wide-area measurement system. *IET Gener. Transm. Distrb.*, 2012, 6(11):1143-1152.

[13] Karapidakis ES, Hatziargyriou ND. Online preventive dynamic security of isolated power systems using decision trees. *IEEE Trans. Power Syst.*, 2002, 17(2):297-304.

[14] Hatziargyriou HD, Contaxis GC, Sideris NC. A decision tree method for on-line steady state security assessment. *IEEE Trans. Power Syst.*, 1994, 9(2):1052-1061.

[15] Guang L, Rovnyak SM. Integral square generator angle index for stability ranking and control. *IEEE Trans. Power Syst.*, 2005, 20(2):926-934.

[16] Gao Q, Rovnyak SM. Decision trees using synchronized phasor measurements for wide-area response-based control. *IEEE Trans. Power Syst*., 2011, 26(2):855-861.

[17] Genc I, Diao R, Vittal V, Kolluri S, Mandal S. Decision tree-based preventive and corrective control applications for dynamic security enhancement in power systems. *IEEE Trans. Power Syst*., 2010, 25(3):1611-1619.

[18] Genc I, Diao R, Vittal V. Computation of transient stability related security regions and generation rescheduling based on decision trees. Presented at: PES General Meeting., Minneapolis, 2010.

[19] Liu C, Sun K, Rather ZH, *et al.* A systematic approach for dynamic security assessment and the corresponding preventive control scheme based on decision trees. Accepted by *IEEE Trans. Power Syst*., 2013.

[20] Senroy N, Heydt GT, Vittal V. Decision tree controlled islanding. *IEEE Trans. Power Syst*., 2006, 21(4):1790-2293.

[21] Diao R, Vittal V, Sun K, *et al*. Decision tree assisted controlled islanding for preventing cascading events. Presented at: IEEE PES Power Systems Conference and Exposition, Seattle, 2009.

[22] Voumvoulakis EM, *et al.* Decision trees for dynamic security assessment and load shedding scheme. Presented at: PES General Meeting, Montreal, 2006.

[23] Rovnyak S, Kretsinger S, Thorp J, *et al*. Decision trees for real-time transient stability prediction. *IEEE Trans. Power Syst.*, 1994, 9(3):1417-1426.

[24] Rovnyak S, Taylor CW, Sheng Y. Decision trees using apparent resistance to detect impending loss of synchronism. *IEEE Trans. Power Delivery*, 2000, 15(4):1157-1162.

[25] Mei K, Rovnyak SM. Response-based decision trees to trigger one-shot stabilizing control. *IEEE Trans. Power Syst.*, 2004, 19(1):531-537.

[26] Guo T, Milanovic JV. On-line prediction of transient stability using decision tree method – sensitivity of accuracy of prediction to different uncertainties. Presented at: PES PowerTech, Grenoble, 2013.

[27] Bernabeu EE, Thorp JS, Centeno V. Methodology for a security/dependability adaptive protection scheme based on data mining. *IEEE Trans. Power Delivery*, 2012, 27(1):104-111.

[28] Figueiredo V, Rodrigues F, Vale Z, Gouveia JB. An electric energy consumer characterization framework based on data mining techniques. *IEEE Trans. Power Syst.*, 2005, 20(2):596-602.

[29] Chien YC, Yuafi HY. Estimation of line flows and bus voltages using decision trees. *IEEE Trans. Power Syst.*, 1994 9(3):1569-1574.

[30] Ugedo A, Lobato E, Peco J, Rouco L. Decision trees applied to the management of voltage constraints in the Spanish market. *IEEE Trans. Power Syst.*, 2005, 20(2):963-972.

[31] Mori H, Kosemura N. Optimal regression tree based rule discovery for short-term load forecasting. Presented at: IEEE PES Winter Meeting, Columbus, 2001.

[32] Mori H, Kosemura N, Ishiguro K, Kondo T. Short-term load forecasting with fuzzy regression tree in power systems. Presented at: IEEE Systems, Man, and Cybernetics Meeting, Tucson, 2001.

[33] Togami M, Abe N, Kitahashi T, Ogawa H. On the application of a machine learning technique to fault diagnosis of power distribution lines. *IEEE Trans. Power Delivery*, 1995, 10(4):1927-1936.

[34] Yong S, Rovnyak SM. Decision tree-based methodology for high impedance fault detection. *IEEE Trans. Power Delivery*, 2004, 19(2):533-536.

[35] Karishnan V, McCalley JD, Henry S, Issad S. Efficient databases generation for decision tree based power system security assessment. *IEEE Trans. Power Syst.*, 2011, 26(4):2319-2327.

[36] Liu C, Rather ZH, Chen Z, *et al*. Importance sampling based decision trees for security assessment and the corresponding preventive control schemes: the Danish case study. Presented at: PES PowerTech, Grenoble, 2013.

[37] He M, Vittal V, Zhang J. Online dynamic security assessment with missing PMU measurement: a data mining approach. *IEEE Trans. Power Syst.*, 2013, 28(2):1969-1977.

[38] Samantary SR. Decision tree-initialized fuzzy rule-based approach for power quality event classification. *IET Gener. Transm. Distrb.*, 2010, 4(4):530-537.

[39] Hong YT, Wen CY, Ching HL. Power system distributed online fault section estimation using decision tree based neural nets approach. *IEEE Trans. Power Delivery*, 1995, 10(1):540-546.

[40] Wehenkel L, Houben I, Pavella M. Automatic learning approaches for online transient stability preventive control of the Hydro-Québec system—part II: a toolbox combining decision trees with neural nets and nearest neighbor classifiers optimized by genetic algorithms. In: *Proc. the 2nd IFAC Symp. Control of Power Plants and Power Systems*, 1995:237–242.

[41] Breiman L. Random forest. *Mach. Learn.* 2001, 45(1):5-32.

[42] Samantary SR, Kamwa I, Joos G. Ensemble decision trees for phasor measurement unit-based wide-are security assessment in the operations time frame. *IET Gener. Transm. Distrb.*, 2010, 4(2):1334-1348.

[43] Kamwa I, Samantary SR, Joos G. Catastrophe predictors from ensemble decision-tree learning of wide-area severity indices. *IEEE Trans. Smart Grid*, 2010, 1(2):144-158.

[44] Kamwa I, Samantary SR, Joos G. On the accuracy versus transparency trade-off of data-mining models for fast-response PMU-based catastrophe predictors. *IEEE Trans. Smart Grid*, 2012, 3(1):152-161.

[45] Friedman JH. Stochastic gradient boosting. *Comput. Stat. and Data An.*, 2001, 29(5):1189-1232.

[46] Salford Predictive Modeler. [Online]. Available: http://www.salford-systems.com/en/products/spm

[47] IBM. IBM SSPS Modeler 14.2 Modeling Nodes. 2011.

[48] Mierswa I, *et al*. Rapid prototyping for complex data mining tasks. Presented at: the 12th ACM SIGKDD Int. Conf. KDDM, 2006.

[49] Bottou L, Cun YL. *Lush Reference Manual*. 2002.

[50] Gawande K, Webers C, Smola A, *et al*. ELEFANT User Manual. Technical Report, NICTA, 2007.

[51] Sonnenburg S, *et al*., Large scale multiple kernel learning. *J. Mac. Learn. Res.*, Mar. 2002, 7:1531-1565.

[52] Spider. [Online]. Available: http://www.kyb.tuebingen.mpg.de/bs/people/spider

[53] Plearn. [Online]. Available: http://plearn.berlios.de/

[54] PyML. [Online]. Available: http://pyml.sourceforge.net

[55] Witten IH, *et al*. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2nd edition, 2005.