Steady State Contingency analysis of electrical networks using machine learning techniques

Dimitrios Semitekos and Nikolaos Avouris

Electrical and Computer Engineering Dep. HCI Group
University of Patras, 26 500 Rio Patras, Greece.
dsem@ee.upatras.gr, avouris@upatras.gr

Abstract. Steady state contingency analysis aims at the assessment of the risk certain contingencies may pose to an electrical network. This is a particularly important task of network operators, especially as network stability issues become of prime importance in the current era of electricity deregulation. The article focuses on the analysis of experimental data that are produced through operating point simulation, contingency application, machine- learning cross validation (based on pre-contingency network index selection algorithms) to point out the "nature" of given contingencies. Experimental statistical results of contingency prediction and selected network state indicators are translated to electric network data in an effort to further interpret the "nature" of each contingency and produce effective predicting algorithms that support operators.

Introduction

Application of machine learning techniques to security assessment of electrical networks has been proposed by many researchers [1]. In this paper we discuss the process of simulating many different operating points (OPs) of a network through the variation of three parameters: the load level, the unit commitment and the network topology [2]. The idea of training of machine learning algorithm on a per contingency basis for subsequent contingency predictions using a preclassified learning set (LS) of operating states or operating points (OPs) is described in detail in [3].

In this paper, a more extended contingency analysis study is presented, where the contingencies examined can be classified in groups of interest. Experimental data demonstrate that combinations of contingencies tend to produce behavior of incremental nature in what concerns their predictability when machine learning tools are used. It can also be assumed that the consequences a contingency may entail to a network may be reflected on the information value of the active or reactive power network indices calculated at pre-contingency times. Sensitivity analysis results about various machine learning algorithms implemented as well as various training to testing split quotas, also lead to interesting results, showing a per contingency consistency in predictions. Experimental results also show that for every contingency to a certain extent the predictive powers of the machine learning tools are influenced by the distribution of the prediction classes in the training data set.

The experimental environment

For the needs of our study the electric network of the Greek island of Crete has been used. fig. 1(a).
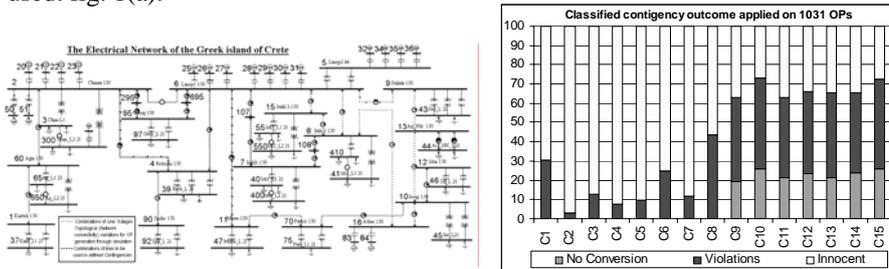

Fig. 1. (a) The electrical network of Crete. (b) The effect of the 15 contingencies studied on the 1031 Operating Points of the network of Crete.

Screened simulated variations in network connectivity, load level and generation plan lead to diversified Operating Points (*OPs)* that are representative of possible states of the electric network. The experimental environment of the simulated *OPs* considered 22 scenarios of line outages. 10 load levels were considered during *OP* simulation ranging from full load scenario to 70% of it. Also 6 different generation scenarios were applied. Thus 22x10x6=1320 *OP* scenarios were simulated, 1031 of which were violations free and were saved in the *OP* repository. Subsequently, 15 line outage events, defined as contingencies under study, were applied on the 1031 simulated *OPs*.

The effect of contingencies on the *OPs* was classified in three discreet categories "innocent", "non-conversion" and "violations". The classification of the contingency outcome for all *OPs* is depicted in fig. 1(b), which demonstrates that there are potentially dangerous (C_8 to C_{15}) and harmless (C_2, C_4, C_5 C_7) contingencies. The contingency effect distribution to classes varies. The questions set focus on the predictability of the effect of these contingencies in relation to selected network characteristics, automatic learning algorithms used and training to testing data split quotas.

A study on feature selection

For most machine learning algorithms that are insensitive to feature redundancy, the selection of additional features will always lead to improvement (even marginal) of the performance achieved, no matter the computational overhead. For instance C4.5 deals remarkably well with irrelevant and redundant information, which is why feature selection has generally resulted in little if any improvement in its accuracy.

In the paper we assume that feature selection algorithms can be applied on the contingencies we study in a "reverse engineering" way, outlining the electrical behavior of the contingencies. For our study, we combined results out of five different

feature selection algorithms and searching methods, included in [4]: The *CfsSubsetEval* algorithm selects a number of the most prevalent features while the *WrapperSubsetEval*, the *InfoGainAttributeEval* and the *GainRatioAttributeEval* algorithms rank all available features according to their predicting powers.

In our study, from a total of 19 candidate features the most prevalent features of the selection algorithms were split into two sets of importance marked in common with "1" and "2" (the most significant ones – selected by both *BestFirst* and *RandomSearch CfsSubSetEval* algorithms). Corollary, the eight most significant features were selected using these algorithms and marked with "1" and "2" (the most significant features). The results from all algorithms were marked in an aggregate table on a scale 1 to 4.

Typical examples of the features of this table (that refer to pre - contingency application saved *OPs* network indices) are:

- *PMargin1,2,4*: Active power margins. Calculated over $PI = \sum_{i=1}^{L} W_i \left(\dfrac{P_i}{\hat{P}_i} \right)^{2n}$, for

  values of *n* equal to 0.5, 1 and 2. [3], [5], [6], [7], [8]

- *QMargin1,2,4*: Reactive power margins. Calculated over $QI = \sum_{i=1}^{NG} W_{Q_i} \left(\dfrac{Q_i}{Q_i^{Max}} \right)^{2n}$,

  for values of *n* equal to 0.5, 1 and 2 respectively [9], [4]. An extensive discussion of indices that can be used for this problem is included in [10].

Table 1. Results from all features selections algorithms used marked in a scale of 1 to 4 (the most significant ones)

Features	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
RedLoad%															
Lines															
TotalRating															
TotalPLoad															
TotalPGen		1													
TotalQLoad		1													
TotalQGen	1	3	1		2	3	2	2	1	1		1		1	2
VoltStabIdx		3	1	1	4	2	4		1				1	1	
MVAFlow			4	3		3	3	4	4	4	4	4	4	4	4
MVAFlDivRat	3	1			1	2		2	4	4	4	4	4	4	4
P%Losses															
PVIndex		2		2	3		2		2	1	2	2	2	2	
QVIndex		3			3										
PMargin1	1		2	2	1	2		2	4	3	3	3	2	3	2
PMargin2	3	2	3	3		1	4	4	1	1	1	1	1	1	1
PMargin4	4	1	2	4	3	4	1	4	1	2	2	3	3	3	4
QMargin1	4	3	1	1		4	1	1		2	2		1		
QMargin2	2		2	2	1		1	1	1		1	1	1		1
QMargin4	4	1	4	4	4	2	4	1	4	3	3	3	4	3	1

This way, for every contingency, four significant sets of features are formed. Features marked with "*4*" named "*red*" areas (the most significant ones), with "*3*" - "*orange*" areas, with "*2*" - "*green*" areas and with "*1*", "*grey*" areas (the less significant ones). In our study, the following sets of features were considered for cross

validation contingency analysis:

- "*Red*" features.
- "*Red*" and "*orange*" features.
- "*Red*", "*orange*" and "*green*" features.
- "*Red*", "*orange*", "*green*" and "*grey*" features.
- *All* available features.

So, our proposed strategy consists of experimenting with the "most significant" sets of features, gradually enriching them with the "less and less significant" ones, ending up using all available features, no matter the redundancy. The proposed technique is useful for assessing the trade off of the enlarging set of features to the computational overhead. Like this, a "threshold" of features can be set for every examined contingency.

Experimental results

For all planned experiments and contingencies applied on the simulated *OPs*, the data were split in 70% - 30 % machine learning training and testing sets respectively and 40-fold cross validations were carried out, considered enough for avoiding bias.

The following machine learning algorithms were implemented: *J48* (C4.5 Decision Trees) [11], *NNge* (Nearest Neighbor with Generalization) [12], *BayesNet* (K2, B Bayes networks) [13], *MLP - MultilayerPerceptron* [14]. The main conclusions of the study are:

- The *NNge* algorithm for all chromatic areas achieves the best predictions and exhibits the least diversions per chromatic area. Best predictions are then attained by the *J48*, the *MLP* and the *BayesNet* algorithms. (fig. 2(b)) After *NNge*, the *BayesNet* and the *J48* algorithms also exhibit smaller diversions per chromatic area than the *MLP* algorithm. (fig. 2(a))
- Fig. 5(b) shows that for the *BayesNet*, the *J48* and the *NNge* algorithms the orange area can be considered as a significant *threshold* set of features combining a relatively high predictability with a rather restricted set of features.

  The *MLP* algorithm exhibits a rather "greedy" behavior in what concerns extending chromatic areas of features (fig. 2(b)).

  On the contrary, the *BayesNet* algorithm for the orange area reaches its highest predictability further addition of features deteriorates its predictability (especially when all features are used). Langley and Sage [15], elucidate this behavior of the Bayesian algorithm mentioning that the predictability of the Bayesian algorithm is improved when redundant features are removed. This property of the Bayesian algorithm shown in fig. 2(b), supports our assumption of considering the orange area of features as a *threshold* set of features.

- It was also found that on a per contingency basis, for all chromatic areas and machine learning algorithms applied, the predictability of each contingency, when compared to the predictability of another contingency, it tends to fluctuate in a similar way. Thus, contingencies C_1 to C_8 tend to exhibit higher prediction rates and

4

lower percentile correct prediction spread rates, while this does not seem the case for contingencies C_9 to C_{15}.
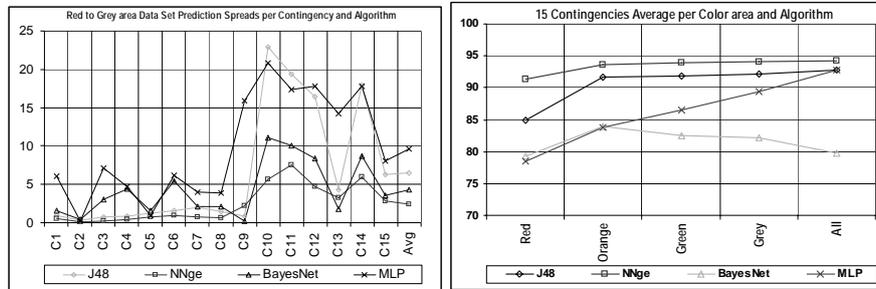

Fig. 2. (a) Maximal averaged percentile correct prediction spread rates among all chromatic areas for the *ML* algorithms (b) 15 Contingency *averaged* percentile correct prediction rates per chromatic area of data (including *all* available features) for the *ML* algorithms

We assume that these permanent fluctuations in the predictability of all contingencies depend on the "*nature*" of each contingency. Attempting to further interpret the role of the "nature" of a contingency, we presume that it can be explained in statistical and electrical terms.

In statistical terms, we suppose that the distribution of the outcome of the application of a contingency for all 1031 OPs to the classes ("Innocent", "Operating violations" and "Non conversion of the power flow algorithm") is related to the predictive powers of the machine learning algorithms. This relation is investigated next.

For this reason, we calculate the standard deviations of the class distribution outcomes of each contingency (fig. 1(b)). Provided that the standard deviations of all contingencies fluctuate between the values 93.11 to 566.91, dividing them by 6, we transform them in a percentile scale. Considering the best prediction curves for the *NNge* and the *J48* algorithms, we also transform the data on a percentile scale using the transformations *J48'(i) = (J48(i) -Min{J48(1) .. J48(15)})*10+2* and *NNge' = (NNge(i)-Min{NNge(1) .. NNge(15)}) *10+5*, where the index *i* refers to any of the 15 contingencies. Then, the average of the *J48'* and *NNge'* prediction is considered. All transformations are depicted in fig. 3.

Comparing the *Std* and the *AVG* curves of fig. 3 for all 15 contingencies, we observe that the initial distribution of the effect of the contingencies to the network is related to the achieved prediction rates of the machine learning algorithms. This can be considered the *statistical* explanation to the *nature* of a contingency. Statistical analysis correlation coefficient ρ was found equal to 0.9605, demonstrating a strong correlation between these two curves.

Discussion of the Results

A rough electric profile of a contingency can be easily outlined through the contingency application outcome class distribution, as well as through the ranking of

the most significant network indices - or machine learning features - that the feature selection algorithms indicate, see Table 1.

This way, fig. 1(b) provides a contingency outcome overview, enabling a general contingency risk assessment; while the view included in Table 1 supports an electric interpretation of the results.
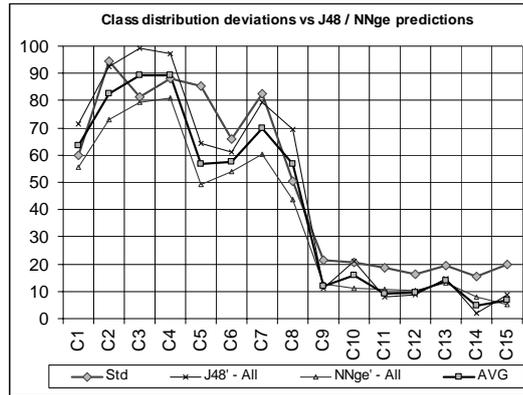

Fig. 3. 15 Contingency linear transformations of the class distribution outcomes standard deviations (fig. 2), versus the transformed *J48'* and *NNge'*, all features predictions.

Indeed, the most significant features can characterize a contingency as potentially harmful to the active power level, or the reactive power level / voltage profile, or both. So, when most significant network indices such as the "*MVAFlow*", "*MVAFlowDivRat*", the "*PVIndex*" or the "*Pmargin1,2,4*", are selected, then the contingency can be considered as potentially harmful at the active power level. When indices such as the "*VoltStabIdx*", the "*QVIndex*", or the "*QMargin1,2,4*" are selected as the most significant ones, then the contingency can be considered as potentially harmful to the reactive power / voltage profile level. When mixed indices are selected, then we can suppose the contingency exhibits a mixed behavior.

If we try to provide an electrical explanation of results such as prediction rates, most significant selected features and contingency distributions, we have to refer to the contingency definitions themselves.

Useful conclusions can come out of such an approach:

Contingencies C_9 to C_{15} include the *MVAFlow* and *MVAFlDivRat* features in the red area (containing the most significant features), while *PMargin1* and *PMargin4* features are included to subsequent orange and green areas. We can so consider these contingencies as potentially harmful to the active power level. From fig. 2, we can also observe similar contingency output distributions. If we refer to the contingency definitions, (fig. 1) we can conclude that the line (6,11) outage, that is common to all C_9 to C_{15} contingencies leads to active power level problems.

Corollary, contingencies like the C2, C5 and C7 indicating more significant features such as the *TotalQGen*, *VoltStabIdx*, *QMargin1* and *Qmargin4* can be considered as contingencies potentially harmful to the reactive power / voltage profile level. Contingencies as C_1 and C_4 exhibit a rather mixed behavior.

Experimental results indicate an incremental effect in what concerns the behavior of certain line outages (contained in certain contingency definitions). Thus, certain contingencies can be characterized by the cumulative effects of the line outages they contain.

For instance contingency C_1, that is a one-line outage contingency - (line (2, 4)) - causes "*Violations*" for 31% of *OP* cases. (fig. 2). On the contrary, the one line outage contingency C_2 - (line (4, 6)) – results in "*Violations*" for less than 4% of the cases.

We can suppose that C_1 is more dangerous than C_2, no matter if both contingencies are neighboring ones, (having bus 4 in common). Indeed, if we refer to the base case scenario, line (2, 4) transmits 100 MW, while line (4,6) 42 MW.

Contingencies C_8 and C_7 are generalizations of C_1 and C_2 respectively, including the (6, 7) line outage too. This is also the case for contingencies C_{10} and C_{11} that are generalizations of C_1 and C_2, including the (6, 11) line outage too.

According to fig. 2, as expected, C_8 is more dangerous than C_7 as well as C_{10} more dangerous than C_{11}. This partially explains the role of the "*nature*" of a contingency that also lies in the role of every outaged line as constituent part of a contingency.

Prediction rates are also lower for more dangerous and in a way more difficult to predict contingencies. So, predictions for C_1 are lower than C_2, as C_8 are lower than C_7, while in what concerns contingencies C_{10} and C_{11}, the differences in predictions are rather small.

Conclusions

Experimental results discussed in this article lead to interesting conclusions:

Graphical visualization of populations of features shows aggregation of outcome classes to Gaussian regions of interest and proves very helpful for contingency analysis. This applies to the distributions of the outcome of contingencies (fig. 2), as well as to the outcome per network index used.

Best *feature selection algorithms* tend to point out to the most significant power transmission indices and / or voltage profile indices, automatically sketching out the nature of a contingency.

The role of the *nature* of a contingency has been investigated. This is partially explained in statistical terms such as the contingency application outcome class distributions, the selected sets of features (chromatic areas) and the machine learning algorithms used. In electrical terms the role of constituent elements of each contingency has been analyzed, such as single line outages contained in contingency definitions, and the incremental effect of multiple components outage.

Experimental results lead to different sets of features that are suitable for predictions of each contingency. A threshold can be set between the number of features and the improved predictions trade-off. Machine learning tools and feature selection tools can be helpful at such a quest, also outlining the expected performance of classes of contingencies from a machine learning perspective.

The predictions have been checked against various training to testing data split quotas and exhibited high prediction rates for all cases, as well as common behaviors on a per contingency basis. Similar behaviors have been identified during data set feature

selection analysis, machine learning tool selection analysis and data set split quota analysis. An optimization of all these factors on a per contingency basis can lead to improved predictions for steady state machine learning contingency analysis.

For the majority of the experiments carried out the nearest neighbor algorithm seems to achieve the best results. Generalizing, we can conclude that the application of instance based learning algorithms to static electric contingency analysis problems presents advantages.

The reported study combined machine learning techniques with statistical modeling, proving that such approaches are suitable for hard engineering applications in such fields like electrical networks analysis and operation.

References

1. Wehenkel L.A., "Automatic Learning Techniques in Power Systems", Kluwer Academic Publ., 1998
2. Cholley P., C. Lebrevelec, S. Vitet, M. de Pasquale, "A Statistical Approach to Assess Voltage Stability Limits", Bulk Power System Dynamics, and Control IV – Restructuring, August 24-28 1998, Santorini, Greece,
3. Hatziargyriou N.D., Contaxis G.C., Sideris N.C., "A decision tree method for on-line steady state security assessment", IEEE Trans. on Power Systems, Vol. 9, Nr. 2, May 1994, p. 1052
4. Ian H. Witten, Eibe Frank, "Practical Machine Learning Tools and Techniques with Java Implementations", Morgan Kaufmann Publishers, 2000
5. Albuyeh F., Bose A., Heath B., "Reactive power considerations in automatic contingency selection", IEEE Trans. on Power Apparatus and Systems, Vol.PAS-101, (1) 1982, p. 107
6. Mikolinnas T.A., Wollenberg B.F., "An advanced contingency selection algorithm", IEEE Transactions on Power Apparatus and Systems, Vol.PAS-100, No.2. February 1981, p.p. 608-611
7. Lauby M.G., Mikolinnas T.A., Reppen N.D., "Contingency selection of branch outages causing voltage problems", IEEE Transactions on Power Apparatus and Systems, Vol.PAS-102, No.12, December 1983, p.p. 3899-3904
8. Yuan-Yih Hsu, Han-Ching Kuo, "Fuzzy-set based contingency ranking", Transactions on Power Systems, Vol. 7, No. 3, August 1992
9. Ejebe G.C., Wollenberg B.F., "Automatic contingency selection", IEEE Transactions on Power Apparatus and Systems, Vol.PAS-98, No.1. Jan/Feb 1979, p.p. 97-109
10. Semitekos D., N. M. Avouris, "A Toolkit for Power systems Security Assessment Based on Hybrid Machine-Learning Techniques", International Journal of Engineering Intelligent Systems, EIS 358, 2004
11. Quinlan J.R., C4.5: Programs for Machine Learning, San Mateo, Morgan Kaufmann, 1993
12. Martin B., "Instance-Based learning: Nearest Neighbor With Generalization", Master Thesis, University of Waikato, Hamilton, New Zealand, 1995
13. Cooper G. and E. Herskovits, "A Bayesian method for the induction of probabilistic networks from data", Machine Learning, 9, 309-347, 1992
14. Minsky M. and S. Papert, "Perceptrons", Cambridge MA, MIT Press, 1969
15. Langley P. and S. Sage, "Induction of selective Bayesian classifiers", Proc. 10th Conference on Uncertainty in Artificial Intelligence, Seattle, W.A, 1994. Morgan Kaufmann.