Electric Power Systems Research 141 (2016) 442–449

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

# Decision tree-based security dispatch application in integrated electric power and natural-gas networks

Denis C.L. Costa*, Marcus V.A. Nunes, João P.A. Vieira, Ubiratan H. Bezerra

*Federal University of Para, Belém, PA, Brazil*

ARTICLE INFO

*Article history:*
Received 9 March 2016
Received in revised form 17 August 2016
Accepted 21 August 2016

*Keywords:*
Integrated energy systems
Natural gas networks
Optimal power flow
Decision tree
Security dispatch

ABSTRACT

This paper proposes a decision tree (DT)-based security dispatch method applied to integrated electric power and natural-gas networks (IPGNs) against credible contingencies that may cause violations. Preventive adjustments to the optimal electric energy generation and gas production are carried out based on the security regions and boundaries of controllable variables determined by the DTs. The easily interpretable DT's rules that describe the security regions are tractable constraints to be included in the optimization routines of electricity generation and gas production rescheduling. Some specific critical contingencies applied to the IEEE 118-bus test system integrated with the 15-node natural gas network are taken as examples to demonstrate a promising application of the proposed security dispatch method to restore IPGN security.



## 1. Introduction

In recent years, the integration of natural gas networks and electric power systems has increased significantly in many countries in the world, due to the growth of gas thermal generation facilities, mainly, the combined cycle plants [1,2]. As a result, the integrated power and natural-gas networks (IPGNs) have become more vulnerable to operational security problems, given the increasing interactions between the gas supply and the electric generators [3].

When a IPGN is vulnerable in a given operating condition due to a probable contingency event, for example, gas leakages in pipelines and/or disconnection of power transmission lines, preventive control actions become indispensable to retrieve the security of both power and gas systems. Dispatching of generating electricity and gas production, with security constraints, is one of the preventive measures that can restore the IPGN from insecure to secure operation state.

Some works in the technical literature present methods of solving problems of security-constrained optimal electric power and gas flow dispatch, applied to IPGN. A formulation of the problem of security-constrained unit commitment with a focus on short-term operation of gas-electricity networks is presented in Ref. [4]. The impact of failures on the gas network upon the electricity market operation is analyzed. However, the model of the gas network is not considered in the formulation. In Ref. [5], the gas network model is integrated to the security-constrained unit commitment model to evaluate the impact of gas and electric power networks interdependence on power system security. However, efforts to restore the power system security are not considered. In Ref. [6], a methodology for solving the security-constrained unit commitment problem with natural gas network constraints is presented. The results show that contingencies in the gas network can impact significantly the security of power systems. However, no preventive action procedure to restore security is proposed in case of happening probable contingencies.

A new formulation of a mixed-integer linear programing security-constrained optimal power dispatch and gas flow is presented in Ref. [7]. For this purpose, a methodology based on the calculation of linear sensitivity factors is proposed to adjust the control variables of the integrated gas and electricity networks, in an optimal and fast way, such that the (N-1) contingencies do not result in security violations. It is worth mentioning that Ref. [7] uses a classic mathematical approach to fit a large set of control variables to restore IPGNs security. On the other hand, the need for coordinated operation among the power systems and gas networks increases the complexity and uncertainty of the security dispatch, once the gas networks are not directly supervised and controlled by power systems operation centers. The advent of automatic machine learning techniques provides a promising approach for setting control variables and their security limits on IPGNs operation.

In recent years, the data mining technique called decision tree (DT) has been widely applied in the area of power systems for the

* Corresponding author.
*E-mail address:* denisclcosta@uol.com.br (D.C.L. Costa).

http://dx.doi.org/10.1016/j.epsr.2016.08.027
0378-7796/

**Nomenclature**

| | |
|---|---|
| B | Bus |
| F | Gas flow |
| NED | Non-electrical demand |
| N | Node |
| P | Active power |
| Pre | Pressure |
| Pro | Production |
| V | Voltage |

purpose of security evaluation and preventive control application [8–10]. Particularly, the DT has as main advantage the interpretability of knowledge learned from the database. The DT differs from traditional techniques for the fact it finds the critical attributes and their thresholds directly from a database using offline simulations. The obtained thresholds not only help to build a predictive model, but also generate security regions for system operators to adopt control actions to ensure the system operation security. Besides that, a DT significantly reduces the set of variables to be used in preventive control actions, allowing operators to stay more focused on the really critical security-related variables of IPGN. Another striking aspect of DT is the fact that it presents a systemic description regarding the critical variables that affect the IPGN security. The systemic character is important, because the set of critical variables, for each IPGN topological configuration, can be distributed by several points of the electrical system and the gas network, often in places that wouldn't necessarily be so evident for the operator.

The main contribution of this paper consists of a dispatching methodology with security constraints applied to IPGNs considering a systemic approach based on optimal load flow and gas flow calculation with security regions defined by DTs rules. In this article, the security regions obtained from DTs are employed to provide guidelines for the decision-making process of dispatching electric power generation and gas production.

This methodology first chooses a few controllable IPGN attributes selected by the offline trained DTs as the only contributors to be tuned during the preventive control process. DTs are then trained to identify the security boundary for the purpose of separating secure cases from insecure ones. The obtained boundary is finally used as a guide to design preventive control strategies.

## 2. Decision tree and security regions

DT for sorting purposes is a supervised learning machine tool to solve problems with high dimensionality data. The fundamental principle is to obtain a predictive model to rank a goal using the attributes that contribute directly to this objective. The DT converts a complex classification process in some "if-then" logical statements, in accordance with the limits of the input attributes or their linear combinations.

For training a DT to have a good performance, it is necessary, first, to build a database consisting of a sufficiently large number of cases. Each case is represented by a goal target (for example, secure or insecure), and attributes, such as active and reactive power generation levels, gas production volumes in wells, etc. The DT is then designed to represent a model for identifying critical attributes that affect the objective target more effectively and directly.

The DT model has a binary structure with two types of nodes, the inner node with two successors and the terminal node without any successor. For each terminal node, also named leaf, a classification result will be assigned in accordance with the goal majority class, as secure or insecure. The classification process begins from the root node and terminates in a terminal node, where the result of sorting is achieved. After the creation of the DT, a cutting process is carried out for the removal of unnecessary nodes and, finally, reducing the DT final size. The DT algorithm used in this paper is known as classification and regression trees (CART) [11].

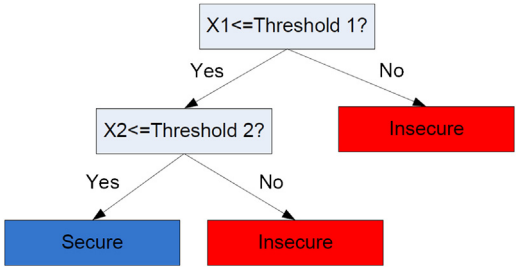


**Fig. 1.** A typical CART.

A typical grown CART looks like Fig. 1, which is resulted from a series of node splits. Given a case represented by set of variables (i.e., x1, x2,. . .) for a particular operating condition, the class (i.e., Secure or Insecure) of the case can be predicted by dropping the variables of the case downward from the root node to a terminal node. A DT is grown by recursive splits of the learning cases at its nodes. The fundamental idea of selecting each split is such that the learning cases in each descendant node are purer than the parent node. The optimal selection of splitting rules can be calculated by repeated attempts to minimize the overall GINI impurity index [11].

The node impurity is maximal when all classes have equal distribution and it is a minimum when there is only one class. Given a data set S, that contains n records, each having a class A, the Gini index of S is given by Eq. (1),

$$\mathrm{Gini}(S) = 1 - \sum_{i=1}^{m} p_i \left[\frac{A}{n}\right]^2 \tag{1}$$

where

$p_i$—is the relative probability of class A in S.

$n$—is the number of records in S.

$m$—is the number of classes.

If S is partitioned into two subsets $S_1$ and $S_2$, one for each link, the Gini index of the partitioned data will be given by Eq. (2),

$$\mathrm{Gini}(S|A) = \frac{n_1}{n}\mathrm{Gini}(S_1) + \frac{n_2}{n}\mathrm{Gini}(S_2) \tag{2}$$

where

$n_1$—is the number of examples of $S_1$.

$n_2$—is the number of examples of $S_2$.

The off-line trained DTs select critical system attributes as good system security indicators and help to build security regions for situational awareness enhancement. The thresholds in the DTs also provide security regions that define operating regions, as shown in Fig. 2. The security region identified by the DT is a simple intersection of the security regions set. With security regions and their boundaries defined by the DT, the security margin can be defined as the smallest distance between the operation point and the security limit.

Two types of DTs can be used to identify the security regions and their contours: the orthogonal and oblique. The orthogonal DTs divide the rules in rectangular regions which are called hyper planes and are orthogonal to the axis in order to associate each region to a class, while the oblique DTs divide the rules that are defined by some linear combinations of attributes and their limits, in non-rectangular regions, dramatically reducing the size of the DT. DTs which cover the partition criterion, divide space $P$ of parameters related to critical attributes. For an orthogonal DT, each