Adaptive Categorization of Complex System Fault Patterns

Seyed Shahrestani
School of Computing and Mathematics
University of Western Sydney
Penrith Campus, Locked Bag 1797
PENRITH SOUTH DC NSW 1797
AUSTRALIA
seyed@computer.org

Abstract: - Due to large amount of information and the inherent intricacy, diagnosis in complex systems is a difficult task. This can be somehow simplified by taking a per-step towards categorizing the system conditions and faults. In this paper, the development and implementation of an approach that establishes class membership conditions, using a labeled training set, is described. More specifically, the use of negative recognition for classification and diagnosis of complex system faults are discussed. The adaptive recognition to achieve the classification is based on discovery of pattern features that make them distinct from objects belonging to different classes. Most of the existing approaches to fault diagnosis, particularly for large or complex systems, depend on heuristic rules. The approach proposed in this work does not resort to any heuristic rules, which makes it more suitable for diagnosis of faults in dynamic and complex systems. For evaluation purposes, using the data provided by the protection simulator of a large power system, its fault diagnosis is carried out. The results of those simulations are also reported. They clearly reveal that even for complex systems, the proposed approach, based on making use of the distinctive features of encountered fault patterns, is capable of fault classification with minimal supervision.

Key-Words: - Adaptive Recognition, Classification, Complex Systems, Diagnosis, Fault Identification, Fault Pattern, Power Systems.

1. Introduction

Power systems have been undergoing an ever increasing complexity to meet the needs of the modern societies. With the higher expectations set for the performance of these systems, their operation, control and management will require operators who are highly qualified and trained. The operators must be able to handle large amounts of data and initiate or modify actions in light of this data. The key factor in this process is the human intelligence.

Knowledge can be considered as the collection of facts, statements whose validities are accepted, and relationships which, when exercised, produce competent performance. Generally, knowledge can be considered to be analytical, heuristic or qualitative [6]. Computations are based on analytical knowledge, heuristics are useful for inference and qualitative reasoning can be based on qualitative knowledge. The many ways that knowledge can be presented and used, give rise to different artificial intelligence methodologies.

Many artificial intelligence based and expert systems have been devised for solving various types of problems in complex systems. Among them, systems for fault identification and diagnosis have received considerable attention, For instance see [3] and [13]. Most of these systems depend on heuristic rules for performing their functions. In most cases, a rule is fired and, if matched, a hypothesis is made in accordance to some hierarchy. The supervisor can then accept or reject the hypothesis. In case of rejection, another rule in accordance with some pre-configured hierarchy is fired that gives rise to a new hypothesis. The process is repeated until either the correct hypothesis is made or all of the appropriate rules have been fired.

But broadly speaking, systems relying on heuristic rules are considered to be brittle. As for such systems, when a new situation falls outside the rules, they are unable to function and new rules have to be generated [5] and [2]. Consequently, a very large knowledge base must be created and stored for retrieving and successful diagnosis purposes. In general, heuristic rules are hard to come up with, and are always incomplete. The rules are usually inconsistent, as in general, no two experts come up with the exactly the same set of rules [6].

This paper describes an approach based on our previous works on classification [6] and [4]. The application of this approach results in formation of fault groups in such a way that distinctive features for all fault locations can be identified. For each fault location a vector, called *mask*, is defined to keep track of the distinctive features of that fault (among a certain group of classes). By making use of the masks and the values of the features signified by them, class identification of a new fault pattern can be carried out. Decision support systems based on this approach can start learning and functioning without resorting to any heuristic rule and their knowledge base is very compact. For evaluation purposes, the proposed approach is tested using data provided by a power system protection simulator. It is shown that, after a thorough training, completely successful and fast identification of all faults can be carried out with minor supervision.

2. AI Applications to Power System

Most of the published work on AI specifically developed for application to power system problems are on building and utilization of expert systems. Among many other areas in power systems, researchers have been actively working on application of AI techniques to solve problems which are hard, if not impossible, to solve by conventional methods. These include problems whose solutions have been traditionally sought through the direct use of human intelligence, i.e. their problem solving loops are closed through human operators. These include diverse application areas, including Energy Management System (EMS), and systems dealing with transient stability providing the ability of the power system to regain steady state stability following a severe disturbance. This is among the areas that AI has found wide acceptance. But the focus of this paper is on fault identification and diagnosis.

Protective relays are used to detect and isolate faults in the shortest possible time. These devices are arranged and coordinated along a network to clear faults in a pre-determined sequence. Their function, isolation of faults, is of prime importance as this can help to reduce the extent of outage and the duration of interruption. Although the sequence of activated breakers can help operators to isolate a fault to some extent, the operation of relays may also result in unnecessary de-energization of part of the network. Fault diagnosis is concerned with identification of faulty parts of the network, so that the steps for re-energizing that part of the network which is not faulty can be taken.

Fault diagnosis is traditionally handled by human operators and therefore is a good candidate for application of AI techniques. Also, there are faults whose identifications are practically possible only by AI approaches, such as high impedance faults. Earlier AI approaches to fault diagnosis comprised of expert systems utilizing shallow knowledge. More recent research has aimed at developing more powerful techniques. The general trend of applications in this area is towards more generality and flexibility for automatic discovery and implementation of concepts, and generic algorithms independent of specific domain areas. Stated alternatively, it is of interest to move towards intelligent identification and control of complex systems. Any meaningful interaction with a system, including its control and fault diagnosis, is not possible without realization of its input/output behavior. Tasks of a planning and forecasting nature, as well as those relating to pattern processing/recognition, can also be thought of as descriptions that relate input information and patterns to output attributes and results.

3. Fault Patterns

In many artificial based systems, pattern recognition finds widespread acceptance for application in diagnosis in complex systems. For instance in power systems, the state of the network exhibits a particular pattern depending on the status of the buses, lines, feeders, links, and breakers which compose it. Each fault can be considered a class and, if the network consists of K_i buses and lines, the goal will be identifying the class of the pattern, or fault location, among the K_i classes. Obviously, patterns representing the same fault location exhibit some common characteristics; i.e have some features and feature values in common, whereas patterns describing a different fault have different values for some or all of these features. In other words, faults can be classified as members of a particular class if they possess some *distinctive features* which make them distinguishable from other fault locations. Consequently, it is logical to form fault groups on the basis of differences, i.e faults which have some evident differences – or distinctive features – from all other faults, are gathered in one group. A feature that may be distinctive for a fault among a particular set of faults is not necessarily distinctive in another set which also includes that particular fault.

In essence, a pattern can be considered as an extract of information regarding various characteristics or features of an object, state of a system, and the like. The pattern of an object with n features under consideration, is normally represented as an n-dimensional vector, $\mathbf{p}_x$. Classification can then be regarded as as the act of partitioning the feature space into K_1 regions or classes, and identification of necessary and sufficient conditions that describe membership criteria for each class, C_x. Clearly, it will be beneficial to have classifiers that can easily accommodate new features and classes. It is also advantageous for these classifiers to be able to process information in parallel, and tune themselves based on their previous experience or misclassifications. That is to develop classifiers which are able to adapt themselves to a new feature space and can partition that space adaptively. Many methods and surveys of these methods for such adaptive pattern recognitions do exist, for example [1] to [3].

The technique to be used here is founded on applying the recognition approaches based upon discovery of distinctive features of each fault [12[7]. These features are those characteristics of a fault which enable one to distinguish it from all others in a group of faults. A feature which is exclusive to a fault – and consequently distinctive – within a group of faults, is not necessarily distinctive of it in another group, which also includes the same fault. Stated in another way, if faults are grouped properly, their distinctive features become evident. Some faults may exhibit exclusive features in the set of all faults, while for others the search must be carried out in a smaller subset.

4. Fault Diagnosis

Operation and control of large dynamical systems is a complex and challenging task. Power systems being indispensable parts of our lives, are among the most complex systems to deal with. The operation of power systems may become even more complex following a fault in some part of the network. Due to this, decision support systems based on artificial intelligence techniques have been of great interest for the management and operation of power systems, particularly as fault diagnosticians.

Pattern recognition is the ability to describe or classify data structures into a set of categories or classes. It is an ability shared by all intelligent beings. To be able to apply pattern recognition techniques to power system fault diagnosis, each fault location can be considered as a class.

Statuses of some of the system components following a fault can then be considered as forming the patterns whose classification is sought. A very important part of this problem consists of proper selection of the components whose statuses must be included in the fault patterns as features that are helpful for diagnosis.

Here we describe an algorithm for establishment of class membership conditions based on making evident the differences among patterns in a labeled training set. The training set may consist of previously encountered patterns whose class membership is known, or it may be formed by the use of an appropriate simulator. Classes in the training set are grouped together in such a way that their exclusive feature values within a group become evident. By making use of these distinctive features and their values, classification of all patterns will be achieved.

The features whose values are found to be distinctive must be included in the patterns; the rest of the features are useless, as far as the classification is concerned.

Fault diagnosis of a typical power distribution system is carried out to test the capabilities of a diagnostician based on the proposed algorithm. Results of extensive tests of different nature show that, after thorough training, fast and successful fault diagnosis is achieved.

Given the complex and large structure of power systems, their fault diagnosis can result in complicated tasks..Various expert and decision support systems have been developed to assist their operators with these tasks, for example see [12]. In a power distribution system, the state of the network depends on the status of its elements; e.g buses, lines, breakers, etc. Each fault can be considered as a particular state of the network. In this way, the status of the elements can be considered as the features that describe fault patterns. This section describes how these features are utilized for identification of fault locations.

Distinctive features for a fault are those whose values are exclusive to it, i.e not repeated in any other fault present in the group. Also, for all patterns representing the same fault, these features must have identical values. For each fault, a *mask* vector whose dimension is equal to the number of features, can be defined; for any distinctive feature, the corresponding element is 1 and for non-distinctive ones the element is 0. With each mask vector, a *mask type* is associated, which is an index to the group of faults which have been used in finding that mask.

If K_1 is the set of all possible fault locations in the training set, then the faults whose distinctive features are evident within the set of all K_1 faults, are associated with type 1 masks. For faults whose distinctive features become evident only with further sub-grouping, higher type masks will be considered.

The mask vector, along with any (single) previously encountered fault pattern, carries the necessary and sufficient conditions for identifying any pattern that may be a consequence of a particular fault. Each pattern is compared with *all* of the other patterns, present in the training set, so that its distinctive features can be identified. These features will then serve as *type one masks* for further class identification purposes. As a classification rule, this means that:

> *Any fault pattern is the consequence of the same fault as the pattern in the training set, if they have the same value for any feature distinguished by the corresponding mask.*

If a mask of type 1 is not associated with each and every fault, higher type masks will be found. To find type 2 masks, the patterns representing faults with type 1 mask will be eliminated from the training set, and the same procedure as the one for finding masks of type 1 is repeated on this new smaller training set. The classification rule will be:

> *If the fault pattern is not due to any fault with a type 1 mask, then it is the consequence of the same fault as the pattern in the training set, if they have the same value for any feature distinguished by the mask (of type 2).*

The whole process is repeated until all faults in the training set have a mask of some type associated with them; in other words classification rules for all faults present in the training set have been found.

During the new fault recognition stage, these rules are fired in the order of mask types. Among the faults with a particular mask type, only one has the possibility of being identified as the fault location for the new pattern. In case that fault location of the new pattern is not identifiable using the highest mask type in the knowledge base, and if its location can be identified by other means (i.e. supervisor), then the pattern will be included in the training set. With this extended knowledge, the masks can be updated and added to the knowledge base.

5. Result Analysis and Discussions

The described approach can be used for development of decision support systems, which can start learning and functioning without resorting to any heuristic rule, while the required knowledge base is very compact. Identification of faults in a power distribution network (Western) is used for the purpose of evaluating the proposed approach. The network consists of 4 feeders (and generators), 4 bus couplers, 35 buses, 35 lines, and 70 breakers. Figure 1 shows one line diagram of the network; More elaborate description of it can be found in [5]. The status of each of the obove, can be taken as a feature. The assigned numerical value of each feature is either 1 (live/closed) or 0 (dead/open). The goal is to identify faulty line or bus; 70 classes. Each pattern will have 148 features. It is worth noting that the status of a bus or a line can actually be deduced from a knowledge of the statuses of the breakers, links, and feeders. That is whether the post-fault patterns contain the lines and buses statuses or not, one should be able to reach proper classification of fault locations; While in the second case, each pattern consists of only 78 features. Such extractions usually cannot be made in expert systems based on heuristic rules. That is if one desired to use a smaller number of features, new (and probably harder to come up with) rules would be needed, and actually the whole system would be changed.

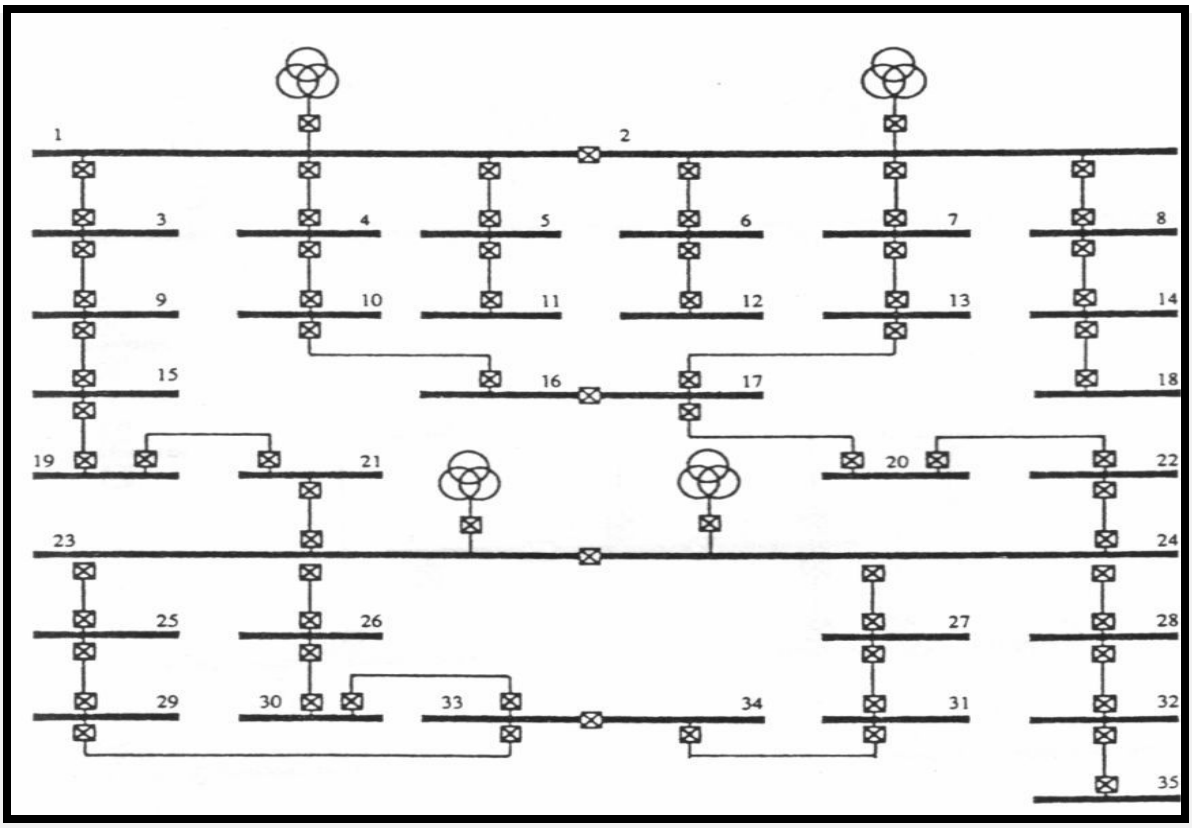

Fig 1 the Western Network

Using a simulator faults were applied to every bus and every line. Ten different cases of faults on two typical power distribution systems are simulated and tested. In each case first a labeled set of faults is presented to the machine to establish the preliminary training; then the capabilities of the machine for diagnosis of other simulated faults are tested. A total of 741 tests that show the effects of noise, effects of knowledge expansion and feature selection capabilities of the algorithm have been devised and carried out. Due to space limitations, the complete description of the tests and test environments cannot be presented here. However, a representative summary is described in this section.

Two cases were considered; First the post-fault statuses of all breakers, links, feeders (and generators), buses, and lines were included in the patterns. In the second case the statuses of buses and lines were excluded from the patterns. In both cases a single pattern for each fault was used as a training set, and masks for all patterns (classes) were found.

In the first case the highest mask type was found to be 4; With type 1 masks mainly corresponding to bus faults. For the second case, the highest mask type was 5; and again most of the type 1 masks were related to bus faults. Table 1 is a summary of mask types, and the number of classes (fault locations) that each mask type has covered for each case.

Table 2 gives the patterns, and resulting masks (in hex) for part of the training set relating to case 2. Fault locations are tagged by bus numbers only, so a fault in bus x is denoted by x,x and a fault in a line between buses x and y is shown as x,y. From table 2 it can be seen that among all classes, only the class corresponding to fault in bus 1 has a 0 for the 75^{th} feature; And only a fault in bus 5 will have a 0 for the value of sixth feature, and so forth. As these features are distinct among *all*

Table 1 Mask types and number of fault locations covered by them

	Number of Classes Covered	
Mask Type	Case 1	Case 2
1	34 Buses, 1 Line	33 Buses, 1 Line
2	29 Lines	1 Bus, 28 Lines
3	1 Bus, 3 Lines	4 Lines
4	2 Lines	2 Lines
5	–	1 Bus

classes, the mask type for their classes will be 1.

For faults in line connecting buses 1 and 4, or the line connecting buses 1 and 5, the interpretation is different as the mask type will not be 1. The masks for faults in these lines are found by eliminating the patterns corresponding to classes whose mask types are 1. For example a fault in the line between buses 1 and 5, is the only fault which can produce a 0 as the value for the fifth feature, among the *classes with mask type 2 or higher*. Consequently, although this feature has the same value for a fault in bus 5 (with a mask type of 1), this would not cease it from being a distinctive feature for a fault in this line. Fault in bus 33 has a mask of FFF … FFF to signify the fact that all of its feature values are distinct, as it is the only class with a mask type 5.

Table 3 gives part of the patterns, and resulting masks (in hex), for the same fault locations as in table 2, but relating to case 1. This table is a continuation of the previous table, in the sense that the statuses of lines and buses are added to previous patterns. Except for the last hex digit, the patterns and masks of case 1, are those shown in table 2 concatenated to the ones in table 3. Fault in bus 33 is a special case, as its mask type is not the same in the two cases, so its complete mask is shown in table 3.

In general, a single distinctive feature is enough for classification purposes; So, comparing the masks in tables 2 and 3, it can be seen that the statuses of the buses and lines are not essential features. For identification of many classes, like classes with mask type 1, these features do not add any useful information. For identification of some other classes they may be considered helpful, but not essential, as it was expected.

Upon completion of the training, tests were made, using data from the simulator. All of the faults could be identified successfully, and there was no need for any mask change. It is clear that, there is no need to train the machine with a complete set at the same time; That is if some patterns (to identify particular classes) are missing during the first training period, the training is not complete, but it is useful in a supervised environment. This was also tested, there were some mask changes, as it should be expected. But again, any class whose presence was established by a pattern

Table 2 Partial training set and the resulting masks for case 2

Fault Location	Pattern	Mask	Mask Type
1, 1	FFFF,FFCF,0FFF,FFFF,FCDC	**0,0,0,0,0020**	**1**
5, 5	F3FF,FFCF,0FFF,FFFF,FFFC	**0400,0,0,0,0**	**1**
2, 8	FFFF,FFFF,0FFF,FFFF,F7FC	**0,0,0,0,0800**	**1**
1, 4	DFFF,FFFF,0FFF,FFFF,FEFC	**2000,0,0,0,0**	**2**
1, 5	F7FF,FFCF,0FFF,FFFF,FCFC	**0800,0,0,0,0**	**2**
33, 33	FFFF,FFFF,0FFF,FFF0,FFBC	**FFF…FFF**	**5**

could be identified successfully. So actually the proposed approach will easily allow any number of classes and patterns to be added gradually, to expand its knowledge base.

6. Conclusions

A method for fast diagnosis in complex systems was described. The described scheme can act as a general decision support system for fault diagnosis in large systems. One of the major advantages of this system is that, the classification process does not rely on heuristic approaches. Its required knowledge base is compact, leading to fast information retrieval and quick and reliable diagnosis The scheme can readily achieve both generalization and specialization of knowledge utilized in the classification process. Updating of the knowledge base and machine learning can easily be implemented. The proposed approach can be of particular interest when the aim is fast classification of a pattern among a large number of classes. The proposed approach can be applied to many classification and fault identification problems in complex systems. The results reported in this paper clearly show that after thorough training, fast and successful identification of all fault patterns in a large network is achieved.

References:

[1] G. Cardoso, J. Rolim, H.Zurn, "Identifying the Primary Fault Section After Contingencies in Bulk Power Systems," *IEEE Trans. Power Delivery*, vol 23, Issue 3, pp. 1335-1342, 2008.

Table 3 Part of the training set and the resulting masks, case 1 (patterns and masks in conjunction with those in table 2)

Fault Location	Pattern	Mask	Mask Type
1, 1	D,1C73,FFFF,8C71,9FFF,F000	**0,0,0,0,0**	**1**
5, 5	F,DF7F,FFFF,EF7D,9FFF,F000	**0,0,0,0,0**	**1**
2, 8	F,FBEE,FFFF,FFEF,1FFF,E000	**0,0,0,0,0**	**1**
1, 4	F,BEFB,FFFF,DEFB,9FFF,F000	**4000,0,2000,0,0**	**2**
1, 5	F,DF7F,FFFF,EF7D,9FFF,F000	**2000,0,1000,0,0**	**2**
33, 33	F,FFFF,FFFD,FFFF,9FFE,7000	**0,0,0,0,0,0,0002,0,0,0** **Complete**	**3**

[2] H. Long, X. Wang, "Aircraft fuel system diagnostic fault detection through expert system," *Proc. IEEE Sixth World Congress on Intelligent Control and Automation*, WCICA2008, pp. 7104-7107, 2008.

[3] G. Mourot, S. Bouschiri, & J. Ragot, "Pattern Recognition for Diagnosis of Technological Systems: A Review", *Proc. of IEEE International Conf. on Systems, Man and Cybernetics*, Vol. 5, pp. 275–281, 1993.

[4] S. Shahrestani, H. Yee, and J. Ypsilantis, "Adaptive Recognition by Specialized Grouping of Classes," in *Proc. 4th IEEE Conference on Control Applications*, Albany, NY, 1995, pp. 637-642.

[5] S. Shahrestani, J. Ypsilantis, and H. Yee, "Adaptive pattern recognition and machine learning of sequences," *Journal of Intelligent Information Processing Systems*, vol. 3, no. 1, pp. 14-22, 1996.

[6] S. Shahrestani, "Exploiting Structure and Concepts for Enhanced Management of Information and Networked Systems," *WSEAS Trans. on Computers*, issue7, vol 4, pp 769-777, 2005.

[7] S. Shahrestani, "Classification in Complex Systems through Negative Recognition," in *Proc. 11th International Conf. on Mathematical Methods, Computational Techniques and Intelligent Systems, MAMECTIS09,* Canary Islands, Spain, 2009.

[8] C. Teo, W. K. Kwok & C. C. Lee, "Application of AI Techniques for fault Diagnosis in Power Distribution", *Proc. IEE International Conf. on Intelligent Systems for Engineering*, pp. 53–58, 1992.

[9] M. Todd, S. McArthur, J. McDonald, S. Shaw, “A Semiautomatic Approach to Deriving Turbine Generator Diagnostic Knowledge,” *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol 37, Issue 5, pp. 979-992, Sept. 2007.

[10] S. Vaucher, H. Sahraoui, J. Vaucher, “Discovering New Change Patterns in Object-Oriented Systems,” *Proc. IEEE International Conf. Reverse Engineering*, pp. 37-41, 2008.

[11] l. Xu, M. Kezunovic, “Implementing Fuzzy Reasoning Petri-Nets for Fault Section Estimation”, *IEEE Trans. Power Delivery*, Vol 23, Issue 2, pp. 676-685, 2008.

[12] Z. Zhang, G. Hope, O. Malik, “Expert Systems in Electrical Power: A Bibliographical Survey", *IEEE Transactions on Power Systems*,Vol. 4, pp. 1355–1362, Oct. 1989.

[13] X. Zheng, Z. Wang, F. Qian, “An Expert System for Real-time Fault Diagnosis and Its Application in PTA Process,” *Proc. IEEE Sixth World Congress on Intelligent Control and Automation*, WCICA 2006, pp. 5623-5627, 2006.

[14] W. Zhao, X. Bai, W. Wang, J. Ding, “A Novel Alarm Processing and Fault Diagnosis Expert System Based on BNF Rules,” *Proc. IEEE International Transmission and Distribution Conf. and Exhibition*, pp. 1-6, 2005.