

# Recognizing Animal-Caused Faults in Power Distribution Systems Using Artificial Neural Networks

Mo-yuen Chow  
IEEE Member

Sue O. Yee  
IEEE Student Member

Leroy S. Taylor  
IEEE Member

Department of Electrical and Computer Engineering  
North Carolina State University  
Raleigh, NC 27695

Distribution Engineering  
Duke Power Company  
Charlotte, NC 28201

## Abstract

Faults are likely to occur in most power distribution systems. If the causes of the faults are known, specific action can be taken to eliminate the fault sources as soon as possible to avoid unnecessary costs, such as power system down-time cost, that are caused by failing to identify the fault sources. However, experts that can accurately recognize the causes of distribution faults are scarce and the knowledge about the nature of these faults is not easily transferable from person to person. Therefore, *artificial neural networks* are used in this paper to recognize the causes of faults in power distribution systems, based on fault currents information collected for each outage. Actual field data collected by Duke Power Company are used in this paper. The methodology and implementation of artificial neural networks and fuzzy logic for the identification of *animal-caused distribution faults* will be presented. Satisfactory results have been obtained, and the developed methodology can be easily generalized and used to identify other causes of faults in power distribution systems.

**Key Words :** Distribution systems, fault identification, animal-caused distribution faults, artificial neural networks, fuzzy logic, conditional probability

## Introduction

Power distribution systems usually experience many different faults during their operation. The distribution faults defined in this paper are faults that result in equipment damage, and/or trigger the operation of a piece of protective equipment, such as a feeder circuit breaker, primary line fuse, or transformer fuse. Sometimes, distribution faults caused by different factors may yield the same consequences, e.g., short-circuited distribution lines due to disturbances by animals and distribution faults caused by traffic accidents both trip the relay [1-4]. Outage costs for both the utility and the customer can be considerable due to distribution faults. If the causes of the faults are known,

then appropriate actions can be taken early to reduce the cost of reparation and to increase the security of the power system.

Each year at Duke Power Company, there are approximately 20,000 distribution faults that open breakers and line fuses. Most utilities rely on general maintenance programs in an effort to reduce fault sources, but these programs can be both costly and ineffective in combating distribution faults if the specific fault causes remain unknown. It is a difficult task to correctly recognize the causes of faults in the distribution system based on the available but limited information [4]. The identification of fault causes is important in the realities of real-time day-to-day utility operation. For example, if more than one outage occurs at the same time in different places, then being able to identify the fault causes based on the outage information can help people at Duke Power decide where to dispatch their service crew so as to take care of the more severe and/or hazardous faults before investigating/repairing other non-severe faults. Fault causes investigation requires detailed analysis, as well as years of technological knowledge and experience. The investigator may have to analyze factors such as geographical location, time of day, type of relays tripped, etc., for each distribution fault to determine the causes of distribution faults.

Nevertheless, an *expert* can quickly pin-point the cause and probably the location of the fault in the distribution network based on the available information. Once the correct cause(s) are identified, specific actions can be taken to eliminate the fault sources. As circuits vary in design vintage and are subject to differing environmental factors, there are also differing fault sources. It is important to recognize the difference between various fault sources such as poor BIL, failed arresters, animals, tree, etc. Unfortunately, the expertise available to determine the causes of the distribution faults problem is very limited. The cause and effect relationship of the distribution fault current problem is highly nonlinear and complicated, and no mathematical models are currently available for their analysis. Furthermore, factors such as experience and instincts belong to the *fuzzy logic* realm and play important roles in problem solving skills. Usually, an utility company has a limited number of employees who possess this type of knowledge and experience. It is expensive to train and hire such experts in every distribution office, aside

92 SM 505-8 PWRD A paper recommended and approved by the IEEE Transmission and Distribution Committee of the IEEE Power Engineering Society for presentation at the IEEE/PES 1992 Summer Meeting, Seattle, WA, July 12-16, 1992. Manuscript submitted January 21, 1992; made available for printing June 11, 1992.

from the fact that this knowledge is not easily transferable from person to person.

*Artificial neural networks* have been used successfully in the past in areas such as fault detection [5,6], control [7] and pattern recognition [8-10]. For this paper, artificial neural networks are used to identify the causes of distribution system faults, based on information such as geographical location, weather and number of phases affected. Artificial neural networks have parallel processing capability, learning and feature extraction ability, robustness to small perturbations and nonlinear modeling capability, and can embed knowledge and experience (*fuzziness*) in their internal structure. Neural networks can learn from available data, and ambiguities can be tolerated to some degree. Because of these properties, neural networks are useful for identifying the causes of distribution fault currents, since the nature of the problem is highly nonlinear and no mathematical models exist for their analysis.

At Duke Power Company in North Carolina and South Carolina, data are recorded each time a fault current is detected in the distribution system as a result of the operation of a breaker relay, fuse or other protective device. The recorded data contains information such as geographical location, circuit ID number, weather condition and time of the fault occurrence. The twenty-two information items that are recorded for each outage at Duke Power Company are obtained as follows. Some of the information is based on a coding system, designed by Duke Power engineers and summarized in charts, that describe the circumstances under which the outages occurred (e.g., weather condition, season of the year, time of the day). Other information items are based on data collected by equipment that monitor the operation of the distribution system (e.g., protective device, number of phases affected). Given this information, an expert can recognize the causes of the distribution faults around 75 % of the time. It is sometimes critical to know the cause of the faults as soon as possible so that appropriate action can be taken fast and efficiently, saving unnecessary costs, such as power system down-time cost, that are caused by failing to identify the fault sources. Assumptions by non-experts can cost a lot of money because their lack of experience leads them to target the wrong fault causes. In this paper we are specifically interested in identifying *animal-caused distribution faults*, since animals are one of the major fault causes in the Duke Power system.

Because experts that can accurately recognize the causes of fault currents are scarce and the knowledge about the nature of these faults is not easily transferable from person to person, it would be useful to train an artificial neural network to *learn* this knowledge and embed it within its internal structure. Once a neural network model and configuration have been selected and designed, neural networks are rather inexpensive to train and replicate, especially when training data is readily available. Therefore,

artificial neural networks can be used in every distribution office of the utility company, if desired. Also, neural networks can be easily updated and adapted to *model* the present fault behavior.

This paper develops a methodology to implement artificial neural networks for recognizing whether a fault current detected in the distribution system was caused by *animals*, such as squirrels. The methodology involves organizing and analyzing the data collected about the distribution faults, as well as choosing the appropriate variables to input to the *Animal-caused Fault Identification Artificial Neural Network* (AFI-ANN). The results obtained by using artificial neural networks to recognize animal-caused faults has been significant and satisfactory. One immediate use for this methodology will be to determine which faults with unknown causes were actually caused by animals so that appropriate action can be taken. This methodology can be easily extended for the identification of other causes of distribution faults.

### ***Brief Description of Artificial Neural Networks***

For many years, artificial neural networks have been studied with the goal of achieving human-like performance in many fields such as pattern recognition and control [5-10]. These networks are composed of dense interconnections of simple computational elements called *nodes* or *neurons*, which are connected by links with variable *weights*. Artificial neural networks typically provide a greater degree of robustness or fault tolerance because they have many processing units. Moreover, they are able to adapt and learn in time to improve performance based on current results. Therefore, artificial neural networks are useful in applications where no appropriate mathematical models exist for the system under investigation.

*Multi-layer feedforward networks*, which are the most commonly used type of neural networks, contain one or more *layers* of *nodes* between the input and output layers. Information from the *external world* are given to the neural network through the *input layer*. Processing, or *internal representation*, is done in the *hidden layers*, and the results of the computations are given back to the *external world* through the *output layer*. Training a neural network means adjusting the network *weights*, or *internal parameters*, so that the network can model the problem at hand. The most popular *training* or *learning* algorithm is the *backpropagation* algorithm [7-9], which is a generalization of the LMS (least mean-square) algorithm. The neural network is *trained* by initially selecting small random *weights* and then presenting all *training data* repeatedly. Weights are adjusted after every trial using side information specifying the correct class until *weights* converge and a *cost function* (usually the mean-square difference between the *desired* and the *actual* network outputs) is reduced to an acceptable value [8,9]. For a more detailed discussion and

explanation of artificial neural networks and their properties, refer to [7-9].

#### Choice of Neural Network Inputs to Identify Animal-caused Distribution Faults

Twenty-two items of information are recorded for each outage caused by fault currents that occurs in the Duke Power division. One of the twenty-two items of information gathered in the outage data is the CAUSE as reported by the crew restoring service. Of these items, the authors, based on their experience, selected five information items which contain essential information about animal-caused fault currents. These five items are : *circuit ID number*, *weather code*, *time off number*, *phase affected* and *protective device*. The *circuit ID number* uniquely identifies a feeder in the Duke Power distribution system. Presently, there are approximately 1,800 feeders in the system. The *weather code* qualitatively represents the condition of the weather at the time of the outage. There are 9 weather codes, numbered 0 – 8. For example, *weather code 0* represents *fair* weather, while *weather code 2* represents *rainy* conditions. The *time off number* contains the month, day and time (expressed in hours and minutes) on the fault occurrence. *Phase affected* indicates which phases (X,Y,Z) were affected by the fault. Finally, the *protective device* code represents which protective device came on as a result of the fault. Duke Power classifies the protective devices into 9 different categories (coded 0 – 8). These categories include P&T System (which refers to any Primary and Transmission protective device that is located upstream from the main breaker feeding an entire circuit), Circuit OCB and Transformer Fuse.

After choosing the information items that can be used to recognize animal-caused faults, the next step is to find some way to represent (or translate) the information contained in the chosen variables so that it is meaningful to a neural network. These variables could be represented as analog or binary codes, or they could be *translated* into a quantity from which a neural network can easily extract salient features. In this paper, we are specifically interested in recognizing whether a given distribution fault was caused by *animals*, such as squirrels. Future research work could include designing other neural networks that can identify other causes of distribution faults, such as equipment failure or load change.

#### Design of the Animal-caused Fault Identification Artificial Neural Network

With extensive experiments and analysis, the available data about the distribution faults were preprocessed, and values of *conditional probabilities* of the chosen variables are used as inputs to the neural network to identify animal-caused faults. Fig. 1 shows the architecture of the artificial neural network designed to recognize animal-caused faults in power distribution systems (AFI-ANN, or

*Animal-Caused Faults Identification Artificial Neural Network*) based on information collected for each outage.

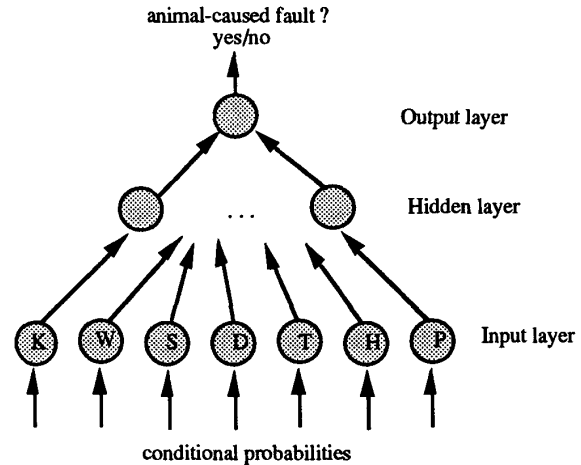


Figure 1. Architecture of AFI-ANN.

Along with the AFI-ANN, there are many data bases that contain the history of previously detected fault currents, sorted by Duke Power sectors such as Chapel Hill and Durham. These data bases contain the faults that have been detected in the past, organized according to the network input variables, which include *circuit ID number* and *weather code*, among others. These data bases are important because the conditional probability values that are used as inputs to the AFI-ANN are calculated from the information contained in these data bases. This means that the conditional probabilities depend on the history of the faults in the distribution system. It seems only obvious to use a priori information (past data) to help recognize the causes of future fault currents.

The designed AFI-ANN has 7 inputs (Fig. 1) that contain valuable and relevant information for the identification of animal-caused distribution faults. Each input unit of the neural network has a symbol associated with it, representing one of the variables shown in Table 1. Note that the information item of *time-off number* is decomposed into *season*, *day of week* and *time of day*.

K	circuit id (depends on geographical location)
W	weather (fair, cold, rain, wind, wind&lightning, lightning, hail, snow, ice)
S	season (spring, summer, fall, winter)
D	day of week (Sunday, Monday, ... , Saturday)
T	time of day (midnight, morning, afternoon, evening)
H	number of phases affected (0 – 3)
P	protective device (P&T system, circuit OCB, line recloser, primary fuse, transformer fuse, transformer CSP, HPP breaker, self clearing, other)

Table 1. List of AFI-ANN input variables.

The output unit of the neural network assumes the values {0.9, 0.1}, depending on whether the given temporary fault current was caused by *animals* or not, respectively (closest neighborhood classification is used when the network output is not exactly 0.9 or 0.1). A flowchart of the development of AFI-ANN is depicted in Fig. 2.

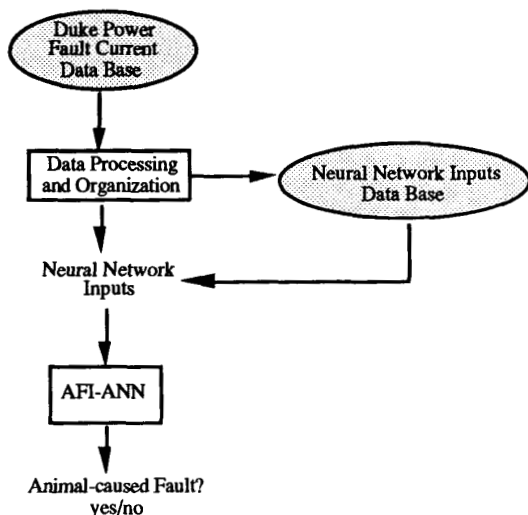


Figure 2. Development of the AFI-ANN.

Duke Power classifies the causes of the fault currents into 12 different categories, coded 0 – 11. *Cause code* 4 represents an animal-caused fault, which is the type of fault studied in this paper. The data used in this paper was obtained from 85 radially fed distribution circuits in the Durham, NC area. These circuits represent approximately 1,300 line miles of three phase, open wye, and single phase overhead feeders and 400 line miles of mostly single phase underground feeders. Approximately one third of this system is operated at 24 KV and two thirds at 12 KV. The system is a multi-grounded wye with a neutral conductor. Single line-to-ground bolted faults range from 300 A to 7,000 A depending on where a fault will occur in the system. Significant portions of urban, sub-urban, and rural areas are included in this line mileage. Fig. 3 shows a histogram of the distribution faults detected in Duke Power's Durham, NC sector during 1989. The data set contains a total of 734 fault cases, with 167 (or 22.75 %) of them caused by *animals* (*cause code* 4). Histograms like the one shown in Fig. 3 are helpful because they give more insight about the fault problem at hand.

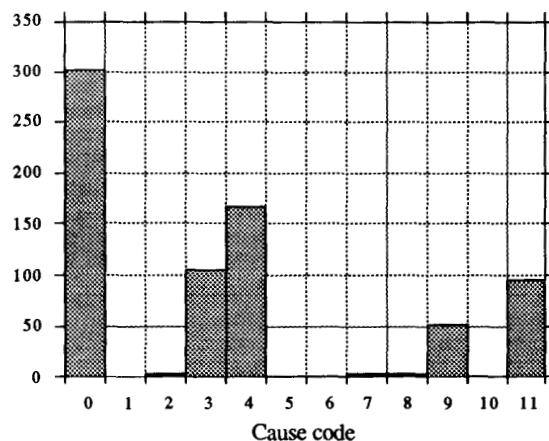


Figure 3. Histogram of the distribution faults detected in Duke Power's Durham, NC sector during 1989.

The *conditional probabilities* capture the statistical characteristics of a particular geographical area. They are given by Bayes' Rule :

$$P(C = c | X = x) = \frac{P(C = c \cap X = x)}{P(X = x)} \quad (1)$$

Eq. (1) represents the probability that the cause of the distribution fault is  $c$  given that event  $x$  is observed, where  $x = [K, W, S, D, T, H, P]$  as shown in Table 1. Using these conditional probabilities helped to keep the AFI-ANN structure simple and easier to implement, since the number of network inputs in this case is 7 — which is expanded from 5 information items of the data base, as discussed in previous sections — instead of 22 inputs if data are taken directly from the Duke Power Fault Currents Data Base (Fig. 2) without preprocessing and are implemented in binary code. By using conditional probabilities as the network inputs, the task of the neural network is to map the nonlinear probabilistic relationship of all the input variables to the output variable.

The authors found the approach of using conditional probability values as inputs to a neural network extremely useful and simple for their research project. This method can be easily extended to other applications of artificial neural networks where the given data has unknown statistical properties but where historical data are available.

With this approach, the neural network is essentially classifying animal-caused distribution faults based on the combination of the different conditional probability values of the given input variables, e.g., weather, season, etc. Since the mapping from the *fuzzy* conditional probability values to the recognition of animal-caused faults is nonlinear and unknown, then neural networks are a suitable tool to implement such mapping.

## Results and Discussion

This paper presents the results obtained using the faults data of Durham collected by Duke Power during 1989-1991. The AFI-ANN was trained using 734 training examples selected from the distribution fault data collected from the Durham, NC area during 1989 (see Fig. 3). AFI-ANN was designed to have 7 input, 15 hidden and 1 output units (7-15-1 network) and was trained using the popular *backpropagation* algorithm, with learning rate  $\eta = 0.3$  and momentum rate  $\alpha = 0.7$  [8,9].

Fig. 4 shows the training accuracy and training error of AFI-ANN. Fig. 4 indicates that the neural network learned the training data (with 99% accuracy) after about 200 training iterations — a training iteration being a single pass through the entire training set. The mean-square training error reached 0.002529 (Fig. 4).

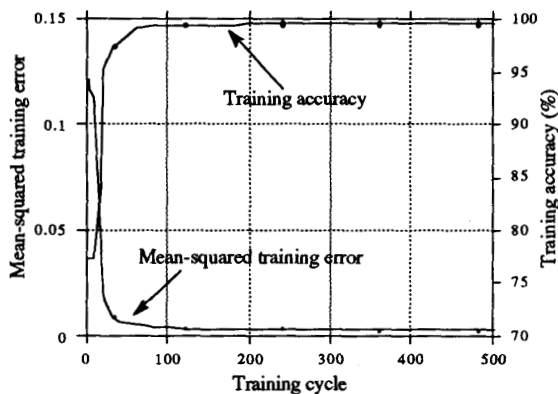


Figure 4. Mean-squared training error and training accuracy of AFI-ANN as a function of training cycles.

Data Set	Accuracy
Training data, Durham 1989 (734 patterns)	99.59 %
Test data, Durham 1990-91 (1337 patterns)	98.57 %

Table 2. Results of AFI-ANN training and testing using Durham, NC data.

The accuracy of the designed neural network is shown in Table 2. The AFI-ANN was trained using Durham, NC data collected during 1989 and tested using Durham, NC data collected during 1990-91. Notice that the AFI-ANN was about 98% accurate in recognizing animal-caused distribution faults previously unseen by the neural network. The 1.43% of the time that the AFI-ANN yields false results, meaning that it predicts animal-caused faults when faults are caused by something other than animals, or it predicts non-animal-caused faults when faults are indeed caused by animals. The overall performance of the AFI-

ANN is highly satisfactory. This means that the neural network is able to predict animal-caused faults based on past data it has seen before. These results show the potential of neural networks as fault identification systems in power systems.

## Conclusions

In this paper, artificial neural networks were used for recognizing animal-caused faults in the Duke Power Company distribution system, using available field data collected for outages. The importance of recognizing the causes of distribution faults was presented. The choice of network input variables and network architecture were explained, and a novel methodology (which uses *conditional probabilities* as network inputs) for training a neural network to recognize animal-caused distribution faults was introduced. The trained neural network is 99 % accurate in recognizing the animal-caused faults of the training data and is above 98 % accurate when it is presented with new and previously unseen data. These results are satisfactory, and the developed methodology can be easily generalized and used to identify other causes of faults in power distribution systems.

## Acknowledgments

The authors of this paper would like to acknowledge the support of Electric Power Research Center project grant 5-31294, North Carolina State University and Duke Power Company. The authors would also like to thank the reviewers of this paper with their valuable comments.

## References

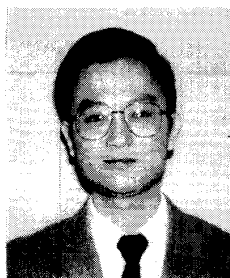
- [1] J.J. Burke, D.A. Douglass, D.J. Lawrence, *Distribution Fault Current Analysis*, EPRI EL-3085, Research Project 1209-1, Final Report, May 1983.
- [2] J.G. Anderson, P.P. Barker, I.S. Grant, D.J. Lawrence, *Characteristics of Lightning Surges on Distribution Lines*, EPRI EL-6782, Research Project 2542-1, Interim Report, May 1990.
- [3] IEEE Tutorial Text, *Detection of Downed Conductors on Utility Distribution Systems*, Course Text 90EH0310-3-PWR.
- [4] Leroy S. Taylor, "Finding the Exact Cause of Fault Currents," Duke Power Company Technical Report, February, 1988.
- [5] Mo-yuen Chow, Sui Oi Yee, "Methodology For On-Line Incipient Fault Detection in Single-Phase Squirrel-Cage Induction Motors Using Artificial Neural Networks," *IEEE Transactions on Energy Conversion*, vol. 6, No. 3, Sept. 1991, pp. 536-545.
- [6] Mo-yuen Chow, Peter M. Mangum, Sui Oi Yee, "A Neural Network Approach to Real-Time Condition

Monitoring of Induction Motors," *IEEE Transactions on Industrial Electronics*, vol. 38, no. 6, Dec. 1991, pp. 448—453.

- [7] Paul J. Werbos, "Backpropagation Through Time : What It Does and How to Do it," *Proceedings of the IEEE*, vol. 78, no. 10, October 1990, pp. 1550—1560.
- [8] Richard P. Lippmann, "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*, April 1987.
- [9] David E. Rumelhart, James L. McClelland, Parallel Distributed Processing, The MIT Press, Cambridge, Massachusetts.
- [10] Yoh-Han Pao, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley Publishing Company, 1989.

### Biography

#### Mo-yuen Chow



Mo-yuen Chow earned the B.S. degree at University of Wisconsin-Madison (June 1982), M. Eng. degree at Cornell University (August 1983), and Ph.D. degree at Cornell University (August 1987), all the degrees are in Electrical Engineering. After graduation, he joined the faculty of North Carolina State University. Currently, Dr. Chow is an Assistant

Professor in the Electrical and Computer Engineering Department at North Carolina State University. He is also the Principal Investigator of various research projects. His research interests include : Power System Analysis, System Monitoring and Fault Detection, Artificial Neural Networks, Fuzzy Logic, Rotating Machines Analysis, System and Control Theory.

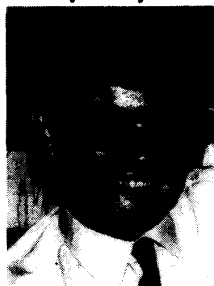
#### Sue Oi Yee



Sui Oi Yee received the B.S. degrees in Computer Engineering (Dec '89) and Electrical Engineering (May '90), and M.S. degree in Electrical Engineering (Dec '91) from North Carolina State University. Her areas of interest include: Neural Networks, Signal Processing and Communication, Rotating Machines, Robotics, Control Theory and

Artificial Intelligence. Ms. Yee is a member of IEEE, Sigma Xi, Phi Kappa Phi, Eta Kappa Nu, Tau Beta Pi and Gamma Beta Phi.

#### Leroy S. Taylor



Leroy S. Taylor is a Senior Distribution Engineer for Duke Power Company, and is a registered professional engineer in North Carolina. He was born in 1949 in Greenville, N.C. and received a B.A. degree (Physics) from the University of North Carolina in 1971. Joining Duke Power in 1977, he acquired extensive experience in distribution system engineering, operation, and construction. Since 1987 he has conducted intensive investigations on the cause of power quality disturbances which originate in the distribution system. He has also redesigned several Duke Power mainframe reporting systems used to evaluate and improve distribution system reliability and power quality.

### Discussion

**Raj Aggarwal** (School of Electrical Engineering, University of Bath, Claverton Down, Bath, Avon BA2 7AY, England): The authors have presented an interesting concept in the application of neural networks for identifying faults caused by animals on distribution systems. A neural network, when fully trained, looks for certain distinct features in the fault data, unique to a particular type of fault (in this case an animal-caused fault). It would be interesting to hear from the authors as to what methodology they adopted for first identifying, and then extracting specific features from animal-caused fault data (as opposed to data from faults such as those caused by lightning). Did they, for example, use simulation techniques for modelling fault transients for faults specifically caused by animals to look for these unique features in the first place?

Manuscript received August 17, 1992.

**M. Y. Chow, S. O. Yee, and L. S. Taylor:** The fact is that the detailed knowledge needed to identify the causes of distribution faults that

occur in power systems is difficult to express using conventional mathematical forms or in a set of rules. But the experts' general knowledge in this area can be easily obtained. For our investigations, we first discussed with several experts and collected a set of factors and guidelines that could be used to identify the causes of distribution faults. We then extracted a set of meaningful variables, preprocessed them and used them as inputs to train a neural network to perform fault cause identification. Our neural network was trained using actual data that was collected and provided by Duke Power Company (Charlotte, North Carolina). We did not choose to use computer simulations to collect the needed data because simulations usually require a mathematical model, and in our application the behavior of the distribution faults is a highly nonlinear and stochastic phenomenon that is *very* difficult to represent accurately through mathematical models. Besides, we already had actual data that was readily available. Once the set of inputs was chosen, the neural network could extract the corresponding physical relationships and certain unique features via network training using the available *real-world* data. This is indeed one of the major advantages of using artificial neural network to identify the causes of distribution faults in power systems.