

“Fault Diagnosis of Power Distribution System using Interval Type-2 Fuzzy Data mining”

¹ Neelam Sahu, ²Manoj Jha, ³ M. F. Qureshi

¹Ph.D. Scholar, Computer Science Department, Dr. C.V. Raman University, Bilaspur, India
(neelam.sahu16@rediffmail.com)

²Department of Applied Mathematics, Rungta Engg. College, Raipur, India
(manojjha.2010@rediffmail.com)

³Department of Electrical Engg., Govt. Polytechnic Dhamtari, India
(mfq_pro@rediffmail.com)

Abstract

In this paper a noble method of fault locating in a radial power distribution system is discussed. Interval type-2 fuzzy system will ease the calculation involved. The history associated with fault occurred in power system can be effectively used as a strong database. This database will act as the trainer to the interval type-2 fuzzy expert system. A rich history of faulty area will enhance the performance of the interval type-2 Fuzzy Fault Identification System (IT2FFIS). Fault section location aims to identify the faulty area in the power system using the post fault status of protective relays and circuit breakers and diagnose the cause of a fault that has occurred. Restoration cost can be significantly reduced by accurately identifying a fault location and its cause, which also reduces the restoration time of the faulty section. In this paper methodology based on interval type-2 fuzzy expert system is discussed for assistance to the operators for easy fault detection. The utility of the soft computing methodology is highlighted. Generally interval type-2 fuzzy sets are suitable for handling the issues related to understandability of patterns, incomplete/noisy data, mixed media information and human interaction, and can provide approximate solutions faster. Neural networks are nonparametric, robust, and exhibit good learning and generalization capabilities in data-rich environments. Genetic algorithms provide efficient search algorithms to select a model, from mixed media data, based on some preference criterion/objective function. Rough sets are suitable for handling different types of uncertainty in data. The interval type-2 fuzzy sets are used here for data mining for fault location in the distribution system.

Keywords: Interval Type-2 Fuzzy Expert System (IT2FES), Power Distribution System, Knowledge Discovery in Database (KDD), Supervisory Control and data acquisition systems (SCADA), Data Mining (DM).

I. Introduction

Data mining is a form of knowledge discovery essential for solving problems in a specific domain. New knowledge may be obtained in the process while eliminating one of the largest costs, viz., data collection. Power system failure data, for example, often exists in vast quantities in an unstructured format. The application of data mining can facilitate systematic analysis in such cases. Soft computing is a consortium of methodologies that works synergistically and provides, in one form or another, flexible information processing capability for handling real-life ambiguous situations. Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve tractability, robustness, and low-cost solutions. Soft computing methodologies (involving fuzzy sets, neural networks, genetic algorithms, and rough sets) are most widely applied in the data mining step of the overall KDD process. The term KDD refers to the overall process of *knowledge discovery in databases*. Fuzzy sets (type-1 & type-2) provide a natural framework for the

process in dealing with uncertainty. Neural networks and rough sets are widely used for classification and rule generation. Genetic algorithms (GAs) are involved in various optimization and search processes, like query optimization and template selection. KDD refers to the overall process of turning low-level data into high-level knowledge. An important step in the KDD process is data mining. Data mining is an interdisciplinary field with a general goal of predicting outcomes and uncovering relationships in data.

II. System Description

The complexity of today’s power system has grown enormously and to have efficient operation of power transmission and distribution network Supervisory Control and data acquisition systems (SCADA) are in use. SCADA is being though widely used in power system but its accuracy has always been a question mark. In Fig.1, zone-1 trips for faults from breaker CB1 to 80 percent of the line length toward breaker CB2. Zone two operates for faults beyond bus Y including some percentage of the transmission line between breakers CB4 and CB3 and the transformer supplying the load at bus Y. Zone three looks backward toward the source and covers the step-up transformer, plus a percentage of the transmission line between breaker CB3 and CB4. Consider a fault on the transmission line between breakers CB3 and CB4.

Assume that breaker CB3 operates correctly but breaker CB4 does not. When a breaker fails to operate, the fault is then removed by back- up breakers. In such cases, the outage range is very large and it is difficult for the dispatchers to judge where the fault is located. Moreover the relay time settings in power system are much smaller and consequently it gets difficult to set the time delays with enough accuracy.

In case of such failures serious implications may arise which may affect both i.e. customer and the power system. Occasional cases may arise where multiple faults may occur, with many breakers being tripped within a short period. Under such situation many alarm messages will be generated and will be received by the dispatch centre which becomes impossible for the dispatchers to analyze the condition with satisfaction and accuracy. The problem can be overcome through the trained system discussed which will employ the interval type-2 fuzzy logic calculations for fault section identification.

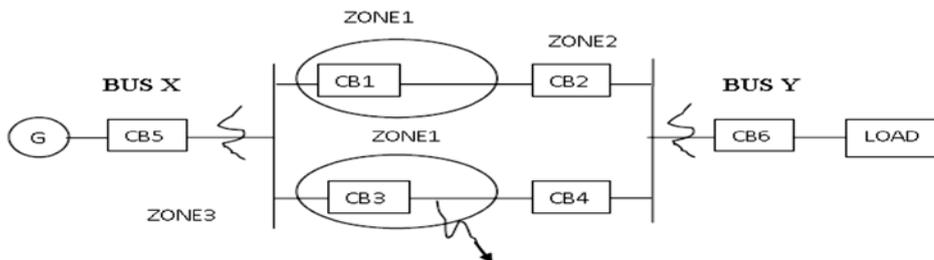


Fig.1 Three step Zone

III. Interval Type-2 Fuzzy Logic System

In this paper a system is discussed for assistance to the operators in fault condition by the help of interval type-2 fuzzy system. This is done by using the post fault status of the circuit breakers and relays. It then calculates membership function for each possible fault section. The objective of this calculation is to determine the likelihood of each candidate fault section as the actual fault section. Moreover membership function provides a convenient means of ranking among possible (or candidate) fault sections, and are the most important factors in decision making. During decision making, the most possible fault section is determined by maximum selection method. In this method most possible fault section is the one which is having highest membership grade. MATLAB code for the proposed scheme is developed and the results obtained in two cases for a power- system network.

Definition of a type-2 fuzzy set and its associated terminology

Type-1 fuzzy logic controllers employ crisp and precise Type-1 fuzzy sets. For example, consider a Type-1 fuzzy set representing the linguistic label of ‘Low’ voltage in Fig.2(a). If the input voltage x is 115v, then the membership of this input to the ‘Low’ set will be the certain and crisp membership value of 0.4. However, the center and end points of this Type-1 fuzzy set can vary due to uncertainties (which could arise for example from noise) in the measurement of voltage (numerical uncertainty) and in the situations in which 115v could be called low (linguistic uncertainty). If this linguistic label was employed with a fuzzy logic controller, then the Type-1 fuzzy logic controller would need to be frequently tuned to handle such uncertainties. Alternatively, one would need to have a group of separate Type-1 sets and Type-1 fuzzy logic controllers, where each fuzzy logic controller will handle a certain situation. On the other hand, a Type-2 fuzzy set is characterized by a fuzzy MF, i.e. the membership value (or membership grade) for each element of this set is itself a fuzzy set in $[0,1]$. For example if the linguistic label of ‘Low’ voltage is represented by a Type-2 fuzzy set as shown in Fig. 2b, then the input x of 115v will no longer have a single value for the MF. Instead, the MF takes on values wherever the vertical line intersects the area shaded in gray. Hence, 115v will have *primary membership* values that lie in the interval $[0.15, 0.65]$.

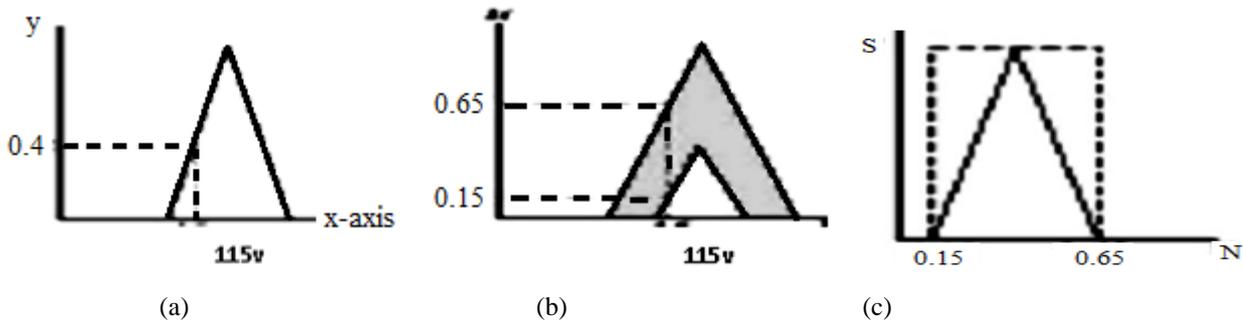


Fig.2 a) A type-1 fuzzy set. b) A type-2 fuzzy set- primary membership function. c) An interval type-2 fuzzy set secondary MF (drawn with dotted lines) and a general type-2 MF (solid line) at a specific point x .

Each point of this interval will have also a weight associated with it. Consequently, this will create an amplitude distribution in the third dimension to form what is called a *secondary membership function*, which can be a triangle as shown in Fig.2c. If the secondary membership function is equal to 1 for all the points in the primary membership and if this is true for $\forall x \in X$, we have the case of an Interval Type-2 fuzzy set (IT2FS). The input x of 115v will now have a primary membership and an associated secondary MF. The MFs of type-2 fuzzy sets are three dimensional and include a Footprint of Uncertainty (FOU) (shaded in gray in Fig.2b). It is the new third-dimension of Type-2 fuzzy sets and the FOU that provide additional degrees of freedom and that make it possible to directly model and handle the numerical uncertainties (for example, uncertainty in video packet delay measurements) and linguistic uncertainties (for example, uncertainty, in modeling cross-traffic characteristics). A Type-2 fuzzy set \tilde{A} is characterized by a T2 MF $\mu_{\tilde{A}}(x, u)$ where $x \in X$ and $u \in J_x \subseteq [0, 1]$, i.e.,

$$\tilde{A} = \{((x, u), \mu_{\tilde{A}}(x, u)), \forall x \in X, \forall u \in J_x \subseteq [0,1]\}$$

in which $0 \leq \mu_{\tilde{A}}(x, u) \leq 1$. \tilde{A} can also be expressed as follows

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x, u) / (x, u), J_x \subseteq [0,1]$$

Where \int denotes union over all admissible x and u . For where \int denotes union over all admissible x and u . For discrete universes of discourse \int is replaced by \cup . A suitable way to describe a T2 MF is in relationship to its secondary MFs. For example, when $f_x(u)=1, \forall u \in J_x \subseteq [0,1]$, then the secondary MFs are interval sets, and if this is true for $\forall x \in X$, we have the case of an *interval type-2 membership function*, which characterizes the Interval Type-2 fuzzy sets (IT2FS) that will be used in this paper. Interval secondary MFs reflect a uniform uncertainty at the primary memberships of x . Since all the memberships in an interval set are unity, an interval set is represented just by its domain interval which can be represented by its left and right end-points as $[l, r]$. The two end-points are associated with two Type1 MFs that are referred to as *Upper MF*

(UMF) and Lower MF (LMF). The UMF and LMF are bounds for the FOU (\tilde{A}) of an IT2 fuzzy set \tilde{A} . The UMF is associated with the upper bound of FOU (\tilde{A}) and is denoted by $\bar{\mu}_{\tilde{A}}(x), \forall x \in X$. The LMF is associated with the lower bound of FOU(\tilde{A}) and is denoted by $\underline{\mu}_{\tilde{A}}(x), \forall x \in X$. The interval type-2 fuzzy set \tilde{A} can be represented in terms of upper and lower membership functions as follows:

$$\tilde{A} = \int_{x \in X} \left[\int_{u \in [\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]} 1/\mu \right] / x$$

This paper's input and output variables will be represented by IT2 fuzzy sets as they are simpler to work and they distribute the uncertainty evenly among all admissible primary memberships. General T2 fuzzy logic control is computationally intensive, whereas the computation simplifies considerably with an IT2 controller using IT2 fuzzy sets. This will enable us to design a T2 FLC that operates in real-time. For discrete universes of discourse X and U, Mendel and Joh have shown that an IT2 fuzzy set \tilde{A} can be represented as follows:

$$\tilde{A} = \sum_{j=1}^n \bar{A}_e^j$$

Where \bar{A}_e^j is an *embedded type-2 fuzzy* which can be written as follows:

$$\bar{A}_e^j = \frac{\sum_{d=1}^N \left[1/u_d^j \right]}{x_d}, u_d^j \in J_{x_d} \subseteq U = [0,1].$$

\bar{A}_e^j has N elements, as it contains exactly one element from $J_{x_1}, J_{x_2}, \dots, J_{x_N}$, namely u_1, u_2, \dots, u_N , each with its associated secondary grade of 1 for IT2 sets. \bar{A}_e^j is embedded in \tilde{A} and there is a total of $n = \sum_{d=1}^N M_d \bar{A}_e^j$, where the M_d are the discretization levels of u_j at each x_d . For discrete universes of discourse X and U, an *embedded type-1 set* \bar{A}_e^j has N elements, as it contains exactly one element from $J_{x_1}, J_{x_2}, \dots, J_{x_N}$, namely u_1, u_2, \dots, u_N , i.e.

$$A_e^j = \sum_{d=1}^N u_d^j / x_d, u_d^j \in J_{x_d} \subseteq U = [0,1]$$

There is a total of $\sum_{d=1}^N M_d A_e^j$. For continuous universes of discourse X and U there is an uncountable number of A_e^j and \bar{A}_e^j embedded within a given type-2 fuzzy set \tilde{A} . Employing T2 fuzzy sets to represent the inputs and outputs of a fuzzy logic controller has many advantages in comparison to T1 fuzzy set.

IV. Structure of Interval Type-2 Fuzzy Expert System (IT2FES)

The system structure will comprise of the elements associated with the distribution system. Basic components of distribution system are transformer, bus-bar, relays, transmission line and circuit breaker. The database consists of the history of alarms received by the SCADA. All the operated circuit breakers and relay's information is stored in the database. As the new alarms are received, they are compared with the existing set of alarms. If the new alarm set matches with the database already in the system, the faulted section is identified easily. If the alarm set is new it will be directed to interval type-2 fuzzy expert system (IT2FES) (see Fig.3) for membership function calculation. Thus a new alarm set will added to the old database with faulty recorded for future reference. Real time data first gets compared with the existing one before going into interval type-2 fuzzy expert system. Thus the fault will be determined at two levels after the data acquisition of all components present in the distribution network. The levels are:

Level 1: At comparison stage if the data matches with the existing faults type in the database.

Level 2: At IT2 fuzzy membership grade calculation stage, in case new set of data is received.

Identification Process

The flow chart shown in Fig.4 gives the steps involved in the process. The process starts with data acquisition. The real time status of the different elements attached in the distribution network is acquired through SCADA system and stored in the database called real time data DB_{new} . According to the data received all the probable fault sections are recorded as FS_i . This data will mainly comprise of number of circuit breakers operated $CB_{operated}$ and number of relays operated $RL_{operated}$. Comparison starts with old database DB_{old} to look for if the case registered is already there or not. Suppose $DB_{new} == DB_{old}$, then fault section is identified and the process ends within no time. Thus the system behaves as a trained module for fast identification of the faulty section. If $DB_{new} \neq DB_{old}$, then data DB_{new} comprising of the $CB_{operated}$ and $RL_{operated}$ is further sent to the interval type-2 fuzzy expert system (IT2FES). The system will calculate interval type-2 membership function for all the relevant faulty sections.

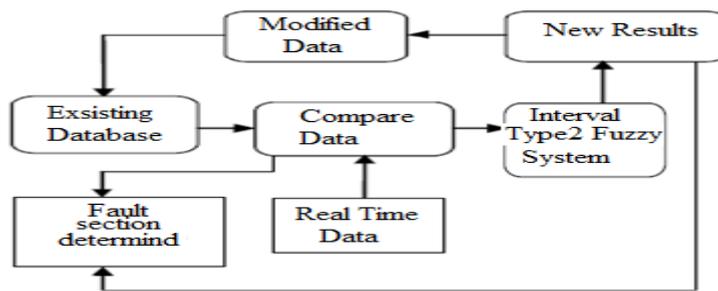


Fig.3 Structure of Interval Type-2 Fuzzy Expert System

The section with maximum interval type-2 membership function will be identified as the faulty section and calculated data will be added to the old database for future reference. Thus the database keeps itself updating with new data and the system gets trained.

Data Handling

Acquiring data is the first step towards fault identification in any distribution network. The data is collected through SCADA system and is sent to diagnostic center for evaluating the data received. After data evaluation decision is taken about faulty section.

Data mining for decision support

Information, such as the section with maximum interval type-2 membership function, a comparison of membership functions among different sections is highly desirable for the faulty section identification. Fig.6 depicts the data mining process for extracting hidden knowledge from large databases. The process focuses on finding interesting patterns that can be interpreted as useful knowledge. It consists of seven steps as shown in Fig.6. This involves the understanding of the interval type-2 fuzzy expert system, its database, and the administrative procedures of the model. A number of mining goals were identified.

The steps involved in data handling for decision support for fault section diagnosis process are: (i) Selection of data, (ii) Data pre-processing, (iii) Data transformation, (iv) Data warehousing, (v) Data mining and (vi) Evaluating the mining results.

The DM and analysis was implemented between the fault information. Authors made sure the causes of the fault and chose the priority solution for troubleshooting by generating candidates IT2 sets and filtering the candidate IT2 set and matching the fault.

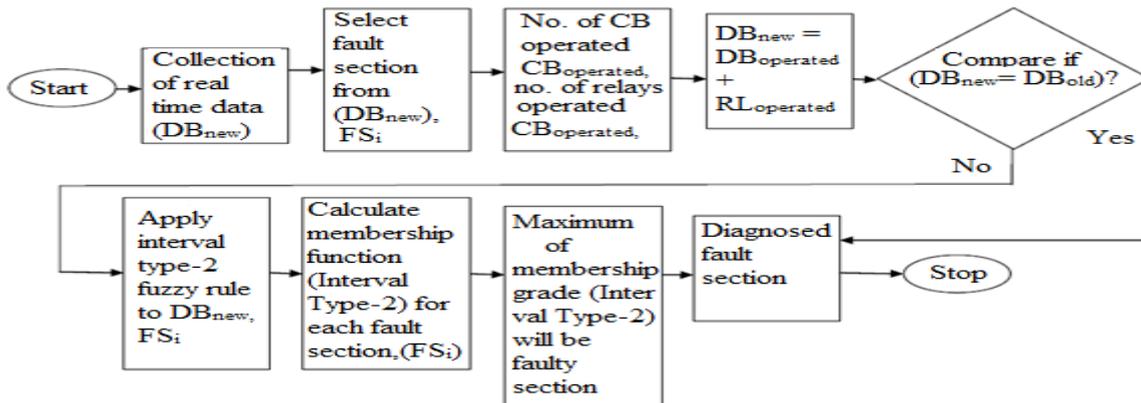


Fig.4 The flow chart showing the steps involved in the fault section diagnosis process.

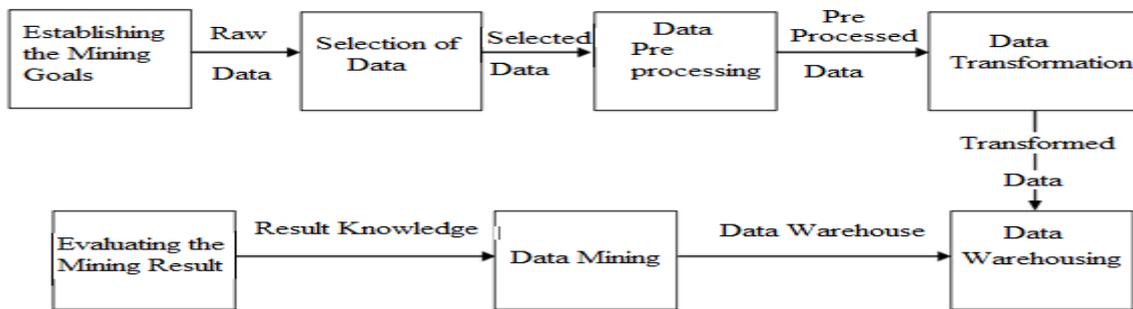


Fig.5 Data mining process for decision support for fault section diagnosis process

DM process in the fault information was shown in Fig.6. The data-mining technique is applied to a power system in NTPC Korba (India) to detect the sections most sensitive to the disturbances. Furthermore, based on the voltage-profile correlation analysis, the influential sections where the most effective voltage adjustment may be strategically applied to assist a sensitive section to recover from the severe voltage fluctuation arising from the disturbance can be deduced.

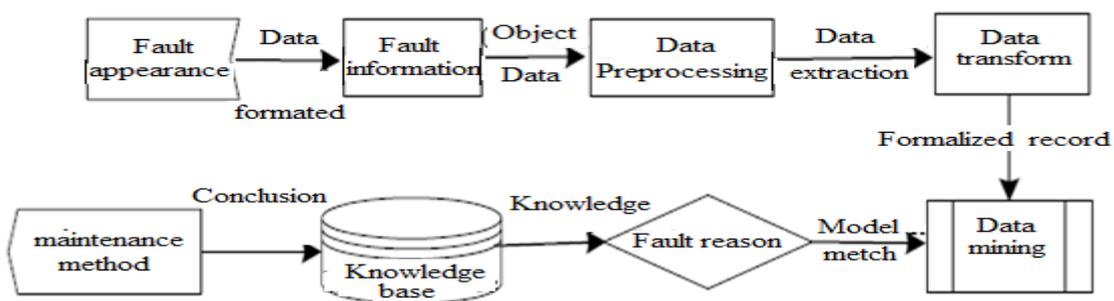


Fig.6. The Data mining process of trouble information.

V. Interval Type-2 Fuzzy Expert System (IT2FES) Simulation:

Referring Fig.6 Structure of Interval Type-2 Fuzzy Expert System that it contains a block i.e. Interval Type-2 Fuzzy Logic System. The simulation of this block is presented here. Mendel and Liang introduced the Type-2 Fuzzy Logic System (T2FLS). Type-2 Fuzzy Logic Systems has a unique characteristic over the conventional type-1 that is the membership function itself became fuzzy-three-dimensional. T2FLS allows the characterization of a type-2 fuzzy set with a superior membership function(Upper) and an inferior membership function(Lower); these two functions can be represented

each one by a type-1 fuzzy set membership function. The interval between these two functions represents the footprint of uncertainty (FOU). The superior membership function is denoted by Upper membership Function (UMF), while the inferior membership function is denoted by Lower Membership function (LMF). Interval type-2 membership function is shown in Fig.7. The structure of a fuzzy logic type-2 system is shown in Fig.8.

Fuzzy logic type-2 system includes 4 components:

***Fuzzifier:** Translates inputs (real values) to fuzzy values.

***Inference System:** Applies a fuzzy reasoning mechanism to obtain a fuzzy output.

***Type Defuzzifier/Reducer:** The defuzzifier translates one output to precise values; the type reducer transforms a Type-2 Fuzzy Set into a Type-1 Fuzzy Set.

***Knowledge Base:** Contains a set of fuzzy rules, and a membership functions set known as the database.

An IT2FLS shown in Fig.8 is also characterized by fuzzy IF-THEN rules, but the membership functions of the ITFLSs are now interval type-2 fuzzy sets (IT2FLSs). From Fig.7, it can be seen that the structure of an IT2FLS is very similar to the structure of a T1FLS and the only difference exists in the output processing block. For a T1FLS, the output processing block only contains a defuzzifier, but for an IT2FLS, the output processing block includes a type-reducer, which maps a T2FS into a T1FS and a defuzzifier.

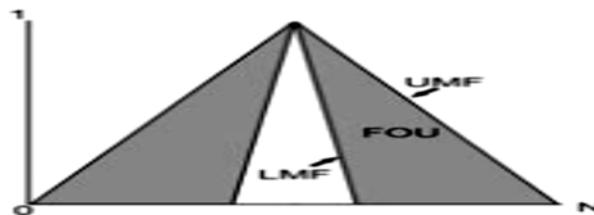


Fig.7: Type-2 Membership Function

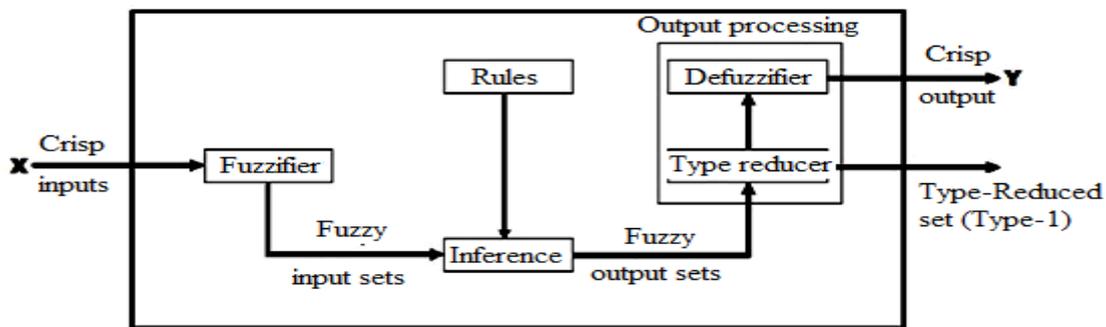


Fig.8 Type-2 Fuzzy Logic System

Assume that there are M rules in the rule base, each of which has the following form:

$$\text{Rule } k : \text{IF } x_1 \text{ is } \tilde{A}_1^k \text{ and } x_2 \text{ is } \tilde{A}_2^k \text{ and } \dots \text{ and } x_p \text{ is } \tilde{A}_p^k, \text{ THEN } y \text{ is } [\underline{w}^k, \bar{w}^k]$$

Where $k = 1, 2, \dots, M$, p is the number of input variables in the antecedent part,

\tilde{A}_i^k k i A ($i = 1, 2, \dots, p, k = 1, 2, \dots, M$) are IT2FLS of the IF-part, and $[\underline{w}^k, \bar{w}^k]$ are the singleton lower and upper weighting factors of the THEN-part. Once a crisp input $X = (x_1, x_2, \dots, x_p)^T$ is applied to the IT2FLS, through the singleton fuzzifier and the inference process, the firing strength of the k th rule which is an interval type-1 set can be obtained as

$$F^k = [\underline{f}^k, \bar{f}^k]$$

$$\text{Where } \underline{f}^k = \underline{\mu}_{\tilde{A}_1^k}(x_1) * \underline{\mu}_{\tilde{A}_2^k}(x_2) * \dots * \underline{\mu}_{\tilde{A}_p^k}(x_p)$$

$$\bar{f}^k = \bar{\mu}_{\tilde{A}_1^k}(x_1) * \bar{\mu}_{\tilde{A}_2^k}(x_2) * \dots * \bar{\mu}_{\tilde{A}_p^k}(x_p)$$

In which $\underline{\mu}$ and $\bar{\mu}$ denote the grades of the lower and upper membership functions of IT2FSs and * denotes minimum or product t-norm. To generate a crisp output, the outputs of the inference engine should be type reduced and then defuzzified.

The type-1 (T1) and Interval Type-2 (IT2) FLC Design

For both Type-1 and Interval Type-2 FLCs triangular MFs were employed as they provide reduced computation time. Choosing the number of MFs is important, as it determines the smoothness of the bit-rate granularity. However, the number of MFs is directly proportional to the computation time and, hence, for this critical power system application, a simplified representation is preferable. In the choice of the number of linguistic labels (or MFs) for both the FLCs input and outputs, the number of linguistic labels were tuned to be the maximum number of labels the system can afford to give a good diagnostic precision, while still enabling the system to achieve real-time performance. Fig.9a describes the MFs for possible fault section $FS_i(CB_{operated})$ which partition its range into four regions which are *Low (L)*, *Medium (M)*, *High (H)* and *Very High (VH)*. Fig.9b describes the MF for possible fault section $FS_i(RL_{operated})$ which has two linguistic labels which are *Increasing Trend (I)* and *Decreasing Trend (D)*. The system output FS_i (Possible fault section i^{th} number) is represented by 11 linguistic labels as shown in Fig.10 which are *Negative Extremely High (NEH)*, *Negative Very High (NVH)*, *Negative High (NH)*, *Negative Medium (NM)*, *Negative Low (NL)*, *Zero (Z)*, *Positive Low (PL)*, *Positive Medium (PM)*, *Positive High (PH)*, *Positive Very High (PVH)*, *Positive Extremely High (PEH)*. The Type1 version of this system employs the center-of sets defuzzification method. The IT2 FLC employs center-of sets type-reduction by using the iterative Karnik-Mendel (KM) procedure. We have employed the iterative KM procedure as the number of used rules is quite small (four rules). Given that the KM procedure is proven to converge in no more than M (number of fired rules, maximum four in our case, if all rules fired) iterations to find the left end point of the type reduced set and no more than M iterations to find the right end point of the type-reduced sets. Hence, the FLC will take a maximum of eight iterations (which is a very acceptable number) to finish the type-reduction step. Thus, it has been feasible to employ the KM procedure and calculate correct values of the type-reduced sets rather than approximate the type-reduced sets using the Wu-Mendel uncertainty bounds method.

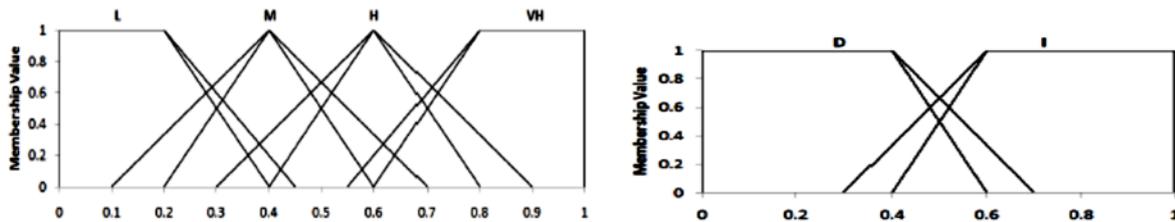


Fig.9. (a) Input Type-2 MFs for $FS_i(CB_{operated})$ (b) Input Type-2 MFs for $FS_i(RL_{operated})$

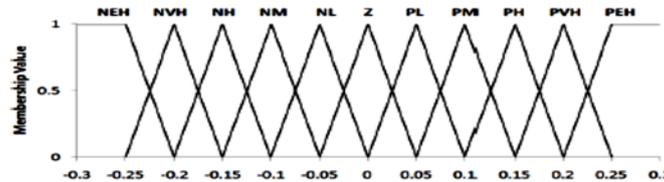


Fig.10. Output Type-2 MFs for FS_i (Possible fault section i^{th} number).

Interval Type-2 FLS input and antecedent operations are shown in Fig.11, it shows that the Min-Pod operations in antecedent signals are same as interval type-1 fuzzy logic system.

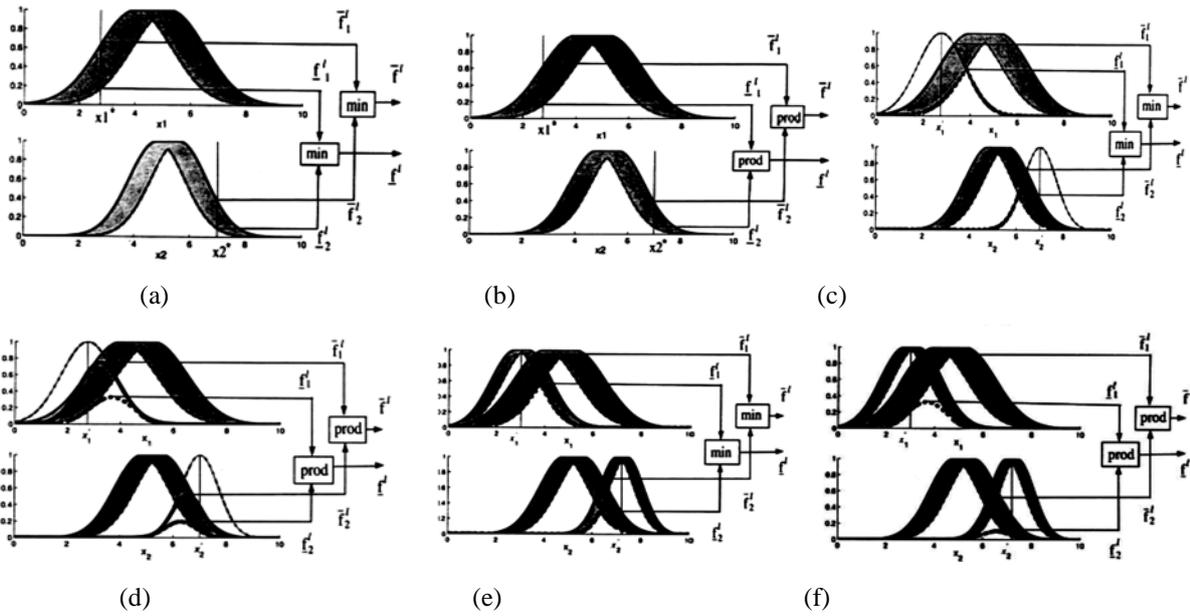


Fig.11 Type-2 FLS: input and antecedent operations.

- (a) Singleton fuzzification with minimum t-norm; (b) singleton fuzzification with product t-norm; (c) NS type-1 fuzzification with minimum t-norm; (d) NS type-1 fuzzification, with product t-norm; (e) NS type-2 fuzzification with minimum t-norm; (f) NS type-2 fuzzification with product t-norm.

The dark shaded regions depict the meet between input and antecedent.

Membership Formulation

The calculation will be determined by the pre fault and post fault status of the circuit breakers and the relays associated with mentioned faulty section. If first step protection has operated then the signals in second and third step protections will be ignored. If first and second step of protection have been isolated as the suspected fault section then the signals of third step protection will be ignored. If all three step protections have not been isolated as the suspected fault section then there is no fault in this particular section. The Upper membership function of FOU is denoted here as ‘U’ in upper case letter and Lower membership function of FOU is denoted here as ‘L’ in upper case letter in the following equations.

$$\begin{aligned}
 FS_i^U, FS_i^L &= [FS_i^U(CB_{operated}), FS_i^L(CB_{operated})] \vee [FS_i^U(RL_{operated}), FS_i^L(RL_{operated})] \\
 &= \{Y_i, \mu_s^U(fault)(Y_i), \mu_s^L(fault)(Y_i) \mid Y_i \in DB_{new}\} \\
 [S_{fault}^U, S_{fault}^L] &= \{FS_i^U, FS_i^L\} \\
 [FS_i^U(CB_{operated}), FS_i^L(CB_{operated})] &= \{Y_i, \mu_s^{UCB}(fault)(FS_i), \mu_s^{LCB}(fault)(FS_i)\} \\
 [FS_i^U(RL_{operated}), FS_i^L(RL_{operated})] &= \{Y_i, \mu_s^{URL}(fault)(FS_i), \mu_s^{LRL}(fault)(FS_i)\}
 \end{aligned}$$

Calculations based on the data of the circuit breakers operated. All the three steps $k= 1, 2 \ \& \ 3$ are considered FS_i^1, FS_i^2 and FS_i^3 membership function is determined.

$$\begin{aligned}
 [FS_i^U(CB_{operated}), FS_i^L(CB_{operated})] &= [FS_i^{U1}(CB_{operated}), FS_i^{L1}(CB_{operated})] \vee [FS_i^{U2}(CB_{operated}), FS_i^{L2}(CB_{operated})] \vee \\
 &[FS_i^{U3}(CB_{operated}), FS_i^{L3}(CB_{operated})] \\
 &= \{Y_i, [\mu_s^{UCB}(fault)(Y_i), \mu_s^{LCB}(fault)(Y_i)]\} \\
 [\mu_x^{UCB}(fault)(Y_i), \mu_x^{LCB}(fault)(Y_i)] &= [\mu_x^{U1}(fault)(Y_i), \mu_x^{L1}(fault)(Y_i)] + [\mu_x^{U2}(fault)(Y_i), \mu_x^{L2}(fault)(Y_i)] + [\mu_x^{U3}(fault)(Y_i), \mu_x^{L3}(fault)(Y_i)]
 \end{aligned}$$

$[\mu_{x(fault)}^{U^k}(Y_i), \mu_{x(fault)}^{L^k}(Y_i)]$ is the interval type-2 membership function to measure the extent to which FS_i belongs to fuzzy set $x(fault)$ by considering only the k^{th} tripped circuit breakers. $[\mu^U(x), \mu^L(x)]$ is the interval type-2 membership function for x (a circuit breaker) to operate correctly. Relays operated in accordance with the circuit breakers are also taken into account. Sometimes relays do operate even though circuit breakers have mal operation. So it is necessary to calculate the membership function of the relays too.

$$[FS_i^{U^1}(RL_{operated}), FS_i^{L^1}(RL_{operated})] = [FS_i^{U^1}(RL_{operated}), FS_i^{L^1}(RL_{operated})] \vee [FS_i^{U^2}(RL_{operated}), FS_i^{L^2}(RL_{operated})] \vee [FS_i^{U^3}(RL_{operated}), FS_i^{L^3}(RL_{operated})]$$

$$[\mu_{x^{RL}(fault)}^{U^1}(Y_i), \mu_{x^{RL}(fault)}^{L^1}(Y_i)] = [\mu_{x^{RL}(fault)}^{U^1}(Y_i), \mu_{x^{RL}(fault)}^{L^1}(Y_i)] + [\mu_{x^{RL}(fault)}^{U^2}(Y_i), \mu_{x^{RL}(fault)}^{L^2}(Y_i)] + [\mu_{x^{RL}(fault)}^{U^3}(Y_i), \mu_{x^{RL}(fault)}^{L^3}(Y_i)] - [\mu_{x^{RL}(fault)}^{U^1}(Y_i), \mu_{x^{RL}(fault)}^{L^1}(Y_i)] \times [\mu_{x^{RL}(fault)}^{U^2}(Y_i), \mu_{x^{RL}(fault)}^{L^2}(Y_i)] - [\mu_{x^{RL}(fault)}^{U^2}(Y_i), \mu_{x^{RL}(fault)}^{L^2}(Y_i)] \times [\mu_{x^{RL}(fault)}^{U^3}(Y_i), \mu_{x^{RL}(fault)}^{L^3}(Y_i)] - [\mu_{x^{RL}(fault)}^{U^3}(Y_i), \mu_{x^{RL}(fault)}^{L^3}(Y_i)] \times [\mu_{x^{RL}(fault)}^{U^1}(Y_i), \mu_{x^{RL}(fault)}^{L^1}(Y_i)] + [\mu_{x^{RL}(fault)}^{U^1}(Y_i), \mu_{x^{RL}(fault)}^{L^1}(Y_i)] \times [\mu_{x^{RL}(fault)}^{U^2}(Y_i), \mu_{x^{RL}(fault)}^{L^2}(Y_i)] + [\mu_{x^{RL}(fault)}^{U^2}(Y_i), \mu_{x^{RL}(fault)}^{L^2}(Y_i)] \times [\mu_{x^{RL}(fault)}^{U^3}(Y_i), \mu_{x^{RL}(fault)}^{L^3}(Y_i)]$$

$[\mu_{x^{RL}(fault)}^{U^1}(Y_i), \mu_{x^{RL}(fault)}^{L^1}(Y_i)]$ is the membership function that Y_i belongs to fuzzy set $x(fault)$ by considering only the operated relays of k^{th} step protection. Thus the data of the membership grades of all the circuit breakers and the relays are tabulated and compared for the results. Decision making is based on two factors. First is the most likely fault section will be the one with the highest grade in comparison with other section's grade. Secondly the section with maximum membership grade i.e. 0.9 will be the faulty section.

VI. Case Studies

The model power system shown in Fig.12. The model consists of 28 sections (L1-L8, T1-T8, A1-A4, B1-B8), 40 circuit breakers (CB1-CB40).

Case Study 1: circuit breakers tripped are CB4, CB5, CB7, CB9, CB12, CB27. In case study 1, the fault area according to the data received through SCADA system is identified as B1, B2, L2, and L4. Now the faulty section identified by data received can vary as fault alarms do get triggered and it is difficult to identify the exact location. In this case actual number of breakers tripped is CB 4, CB 5, CB 7, CB 9, CB 12 and CB 27, whereas the relays operated are 4, 5, 7,9,12, 27, 6, 8, 10. The relays indicating false alarm are relay no. 6, 8 and 10 which operated in secondary zone. Interval Type-2 Membership calculated for the 4 section namely B1, B2, L2, and L4, are [0.866,0.478], [0.845,0.398], [0.724,0.345] and [0.730,0.351] respectively. (See Fig.12 Test network for case study 1 and 2). The type1 membership function after type reduction can be obtained as follows:

$$\mu_{S(fault)}(Y_i) = \frac{\mu_{S(fault)}^U(Y_i) + \mu_{S(fault)}^L(Y_i)}{2}$$

Using above equation the Type1 Membership calculated for the 4 section namely B1, B2, L2, and L4, are 0.672, 0.621, 0.539 and 0.540 respectively. In case of B2, the breakers tripped in primary zone are nil, whereas all the 6 beakers tripped belongs to the secondary zone. Thus the membership calculation will yield less value than the section B1. In case of line L2, only 1 breaker has tripped in primary zone, whereas rest of the 5 breakers have tripped in tertiary zone. Thus again the membership calculated for this section will give less value than B2. Very similarly in case of L4, only CB 27 has operated in primary zone and rest of the 5 CBs have operated in tertiary zone.

The membership grades are now calculated with the help of the formulation discussed in the section 5. Thus less number of rules is there in the calculation, and memory space taken is also less. We also find flexibility in the determination of the membership grades of the sections obtained from SCADA. This system is reliable in terms of the non-sensitivity to unwanted signals.

So Maximum membership calculated belongs to B1 and so this is the most probable faulty section of all the given zones.

Case Study 2: circuit breakers operated are CB19, CB20, CB32, CB33, CB31, CB34, CB35, CB36, CB37, CB39 & CB40. In case study 2, the fault area according to the data received through SCADA system is identified as B7, B8, L5, L6, L7, L8, T7 and T8. Now the faulty section identified by data received can vary as fault alarms do get triggered and it is difficult to identify the exact location. In this case actual number of breakers tripped is CB19, CB20, CB32, CB33, CB34, CB35, CB36, CB37 & CB39. The relays indicating false alarm are relay no. 31 and 40, which operated in secondary zone. Interval Type-2 Membership calculated for the 8 section namely B7, B8, L5, L6, L7, L8, T7 and T8 are [0.822, 0.453], [0.898, 0.498], [0.898, 0.498], [0.754, 0.445], [0.898, 0.498], [0.754, 0.445], [0.898, 0.498], [0.898, 0.498] respectively. Using above equation the Type-1 Membership calculated for the 4 section namely B7, B8, L5, L6, L7, L8, T7 and T8, are 0.632, 0.694, 0.694, 0.599, 0.694, 0.599, 0.694 and 0.694 respectively.

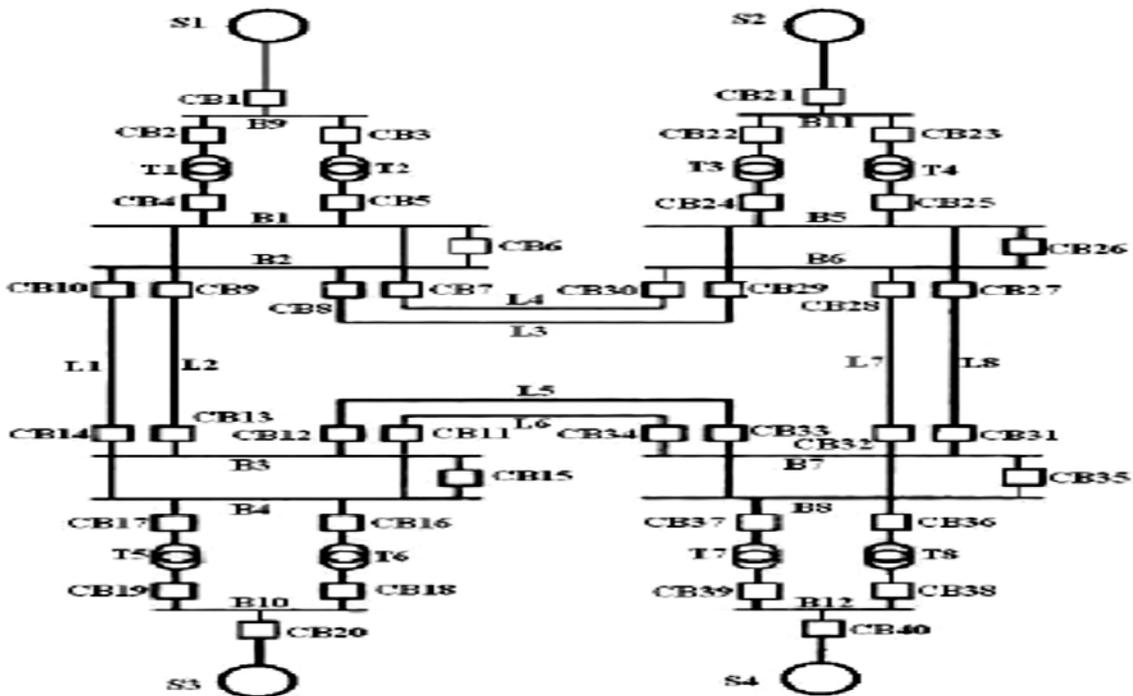


Fig.12 Test Circuit for Case Study

Table 1
Membership Grade for Case Study 1

Possible Fault Sections	Interval Type-2 Membership Grade [$\mu S^U_{(fault)}(Y_i)$, $\mu S^L_{(fault)}(Y_i)$]	Type-1 membership Grades after type Reduction $\mu S_{(fault)}(Y_i)$	Circuit Breakers Tripped
B1	[0.866, 0.478]	0.672	CB4, B5, CB7, CB9, CB12, CB27
B2	[0.845, 0.398]	0.621	CB4, CB5, CB7, CB9, CB12, CB27
L2	[0.724, 0.345]	0.539	CB4, CB5, CB7, CB9, CB12, CB27
L4	[0.730, 0.351]	0.540	CB4, CB5, CB7, CB9, CB12, CB27

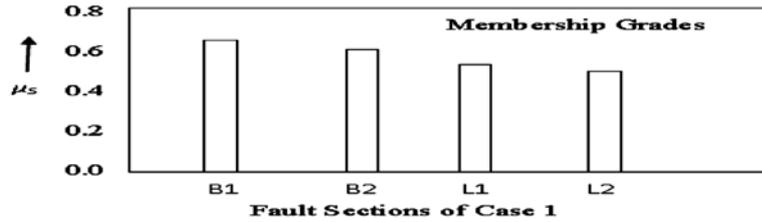


Fig.13 Grades assigned to all sections (Case1)

Table 2

Membership Grade for Case Study 2

Possible Fault Sections	Interval Type-2 Membership Grade [$\mu S^U_{(fault)}(Y_i)$, $\mu S^L_{(fault)}(Y_i)$]	Type-1 membership Grades after type Reduction $\mu S_{(fault)}(Y_i)$	Circuit Breakers Tripped
B7	[0.822, 0.453]	0.632	CB20, CB30, CB33, CB34, CB35
B8	[0.898, 0.498]	0.694	CB32, CB33, CB39
L5	[0.898, 0.498]	0.694	CB19, CB32
L6	[0.754, 0.445]	0.599	CB20, CB30, CB33, CB34, CB35
L7	[0.898, 0.498]	0.694	CB29, CB39
L8	[0.754, 0.445]	0.599	CB20, CB30, CB33, CB34, CB35
T7	[0.898, 0.498]	0.694	CB34, CB36
T8	[0.898, 0.498]	0.694	CB35, CB37

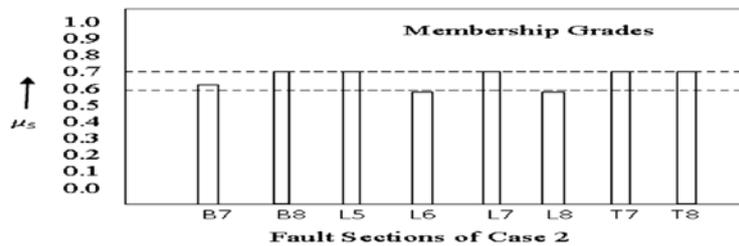


Fig.14 Grades assigned to all sections (Case2)

In this case study the sections with maximum membership value have been identified as B8, L5, L7, T7 and T8. Here as it is clear from the fuzzy calculation the faulty section B7, L6 and L8 are just the part of the section but not the faulted one just because the grades assigned to them is minimum in comparison to rest of the sections. As we compare the data, decision can be quickly made about the faulty section. This calculation will be recorded in the existing database. Thus a database will be constantly updated and can be assessed for the future reference.

VII. Conclusion

This paper describes Fault Section Diagnosis of Power Distribution System using Interval Type-2 Fuzzy Data mining. The trained database compares the real time data with the existing data and gets the decision. For new data set interval type-2 fuzzy system is being used to identify the faulty section. Interval type-2 membership grades calculated for various sections gives the most possible fault section. Three steps are discussed in the interval type-2 membership formulation of the circuit breakers and the relays to get accurate answer. This database will act as the trainer to the interval type-2 fuzzy expert system. A rich history of faulty area will enhance the performance of the Interval Type-2 Fuzzy Fault

Identification System (IT2FFIS). Fault section location aims to identify the faulty area in the power system using the post fault status of protective relays and circuit breakers and diagnose the cause of a fault that has occurred. Restoration cost can be significantly reduced by accurately identifying a fault location and its cause, which also reduces the restoration time of the faulty section. Decision making is based on two factors. First is the most likely fault section will be the one with the highest grade in comparison with other section's grade. Secondly the section with maximum membership grade i.e. 0.7 will be the faulty section.

References

1. L. Xu, M.Y. Chow and L. S. Taylor, (Feb 2007) "Power Distribution Fault Cause Identification with Imbalanced Data Using the Data Mining Based Fuzzy Classification E-Algorithm", IEEE Transaction on Power Systems, Vol.22, No.1, pp.164-171.
2. L. Xu and M.Y. Chow, (2006) "A classification approach for power distribution systems fault cause identification", IEEE Transactions on Power Systems, Vol. 21, No. 1, pp. 53-60.
3. Hong-Chan Chin, (2003) "Fault Section Diagnosis of Power System Using Fuzzy Logic", IEEE Trans Power System, Vol. 18, No. 1, pp. 245-250.
4. Xiaohui Zhul , Xin LuI, Dong Liu, Bingda Zhang, (2010) "An Improved Fault Locating System of Distribution Network based on Fuzzy Identification", China International Conference on Electricity Distribution, pp. 1-6, 13-16 Sept. 2010, Nanjing.
5. Fábio Bertequini leão, Rodrigo A. F. Pereira and José R. S. Mantovani, (2009) "Fault section estimation in electric power systems using an artificial immune system algorithm", International Journal of Innovations in Energy Systems and Power, Vol. 4, No. 1, pp. 14-21.
6. P.K. Dash, H.S. Behera, I.W.C. Lee, (2007) "Time sequence data mining using time–frequency analysis and soft computing techniques", Appl. Soft Comput. J., doi:10.1016/j.asoc.2007.01.001
7. A. Pellicel, et.al. (2011) "Intelligent Data Mining for Turbo-Generator Predictive Maintenance: An Approach in Real-World", Proc. of the International Workshop on Soft Computing Applications and Knowledge Discovery (SCAKD 2011), Moscow, Russia, June 25.
8. He YueShun, and Ding QiuLin. "Fault Mode Analyze of Power System Based on Data Mining". Proc. of the 2009 International Symposium on Web Information Systems and Applications (WISA'09) Nanchang, P. R. China, May 22-24, 2009, pp. 318-320
9. U. M. Fayyad, et. al. "Advances in Knowledge Discovery and Data Mining. Menlo Park, CA: AAAI/MIT Press, 1996.
10. K. J. Cios, W. Pedrycz, and R. Swiniarski, *Data Mining Methods for Knowledge Discovery*. Dordrecht, The Netherlands: Kluwer, 1998.
11. C. K. Shin, S. J. Yu, U. T. Yun, and H. K. Kim, "A hybrid approach of neural network and memory based learning to data mining," *IEEE Trans. Neural Networks*, Vol. 11, pp. 637–646, 2000.
12. N. N. Karnik and J. M. Mendel, "Type-2 fuzzy logic systems: Type-reduction," in *IEEE Syst., Man, Cybern. Conf.*, San Diego, CA, Oct. 1998.
13. N. N. Karnik and J. M. Mendel, "Applications of type-2 fuzzy logic systems to forecasting of time series," *Inform. Sci.*, vol. 120, pp. 89–111, 1999.
14. Satya Prakash and S.C.Gupta, "Fuzzy Logic Based Trained Fault Locating Mechanism in Power Distribution Network" International Journal of Emerging Technology and Advanced Engineering (ISSN 2250-2459), Vol. 2, Issue 7, pp.129-135, July 2012
15. N. N. Karnik and J. M. Mendel, "Type-2 fuzzy logic systems," *IEEE Trans. Fuzzy Syst.*, Vol. 7, pp. 643–658, Dec. 1999.
16. J. M. Mendel and R. I. B. John, "Type-2 fuzzy sets made simple", *IEEE Trans. on Fuzzy Syst.*, Vol. 16, No. 2, pp. 117-127, 2002.
17. H. Shu, Q. Liang and J. Gao, (2008) "Wireless sensor network lifetime analysis using interval type-2 fuzzy logic systems", *IEEE Trans. on Fuzzy Syst.*, Vol. 16, No. 2, pp. 416-427.
18. Q. Liang and J. M. Mendel, "Interval type-2 fuzzy logic systems: Theory and design", *IEEE Trans. on Fuzzy Logic Systems*, Vol. 8, No. 5, pp. 535-550, 2000.
19. D.Wu and W.Tan, (April 2005) "Type-2 FLS Modeling Capability Analysis", *IEEE Int. Conf. on Fuzzy Syst.*, pp. 242-247, Reno, USA.
20. H. Wu and J. M. Mendel, (2002) "Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems", *IEEE Trans. on Fuzzy Syst.*, Vol. 10, No. 5, pp. 622-639.
